Analytics comes of age

McKinsey Analytics

January 2018

McKinsey&Company

About McKinsey Analytics

McKinsey Analytics helps clients achieve better performance through data. We work together with clients to build analytics-driven organizations, providing end-to-end support covering strategy, operations, data science, implementation, and change management. Our engagements range from use-case-specific applications to full-scale analytics transformations. Our teams of consultants, data scientists, and engineers work together with clients to identify opportunities, assess available data, define solutions, establish optimal hosting environments, ingest data, develop cutting-edge algorithms, visualize outputs, and assess impact while building capabilities to sustain and expand it. Learn more at mckinsey.com/business-functions/mckinsey-analytics/our-insights.

Table of contents

Introduction

Since the coining of the term "artificial intelligence" more than a half century ago, AI as a field has experienced a number of hype cycles ending in disappointment. In the meantime, companies successfully deployed other types of analytics on the back of "big data," a term that steadily took over the headlines.

But the buzz around the term "big data" is dwindling. Why? Data is now like air. It's all around us. It has become common knowledge that the world churns out an enormous and expanding amount of data each day—billions of gigabytes, in fact. In industry, all organizations create data, and as storage costs continue to tumble, more of it is being kept and analyzed to create competitive advantages. Increasingly, organizations also share data with other companies to realize new businesses, giving rise to the beginnings of digital ecosystems that stand poised to blur traditional industry borders.

It's this explosion of data that has helped propel AI—particularly machine learning and its subset, deep learning—back into the spotlight. Ubiquitous data provides the fuel that allows these AI models to power a rising number of business applications.

While tech giants are still the biggest investors in AI, incumbents are upping their investments as well. And they should. Machine learning, and increasingly deep learning, are beginning to unlock real value across business functions and in most industries, extending the power of analytics, particularly in organizations with solid digital foundations.

But overall, as analytics comes of age, there are some growing pains. While investments in analytics are booming, many companies aren't seeing the ROI they expected. They struggle to move from employing analytics in a few successful use cases to scaling it across the enterprise, embedding it in organizational culture and everyday decision-making.

Empowering people with analytics—that's where the real value creation occurs. And simply having the best data or writing the most cutting-edge algorithm won't make it happen. It requires a wholesale organizational transformation, complete with robust change management and analytics/AI education programs.

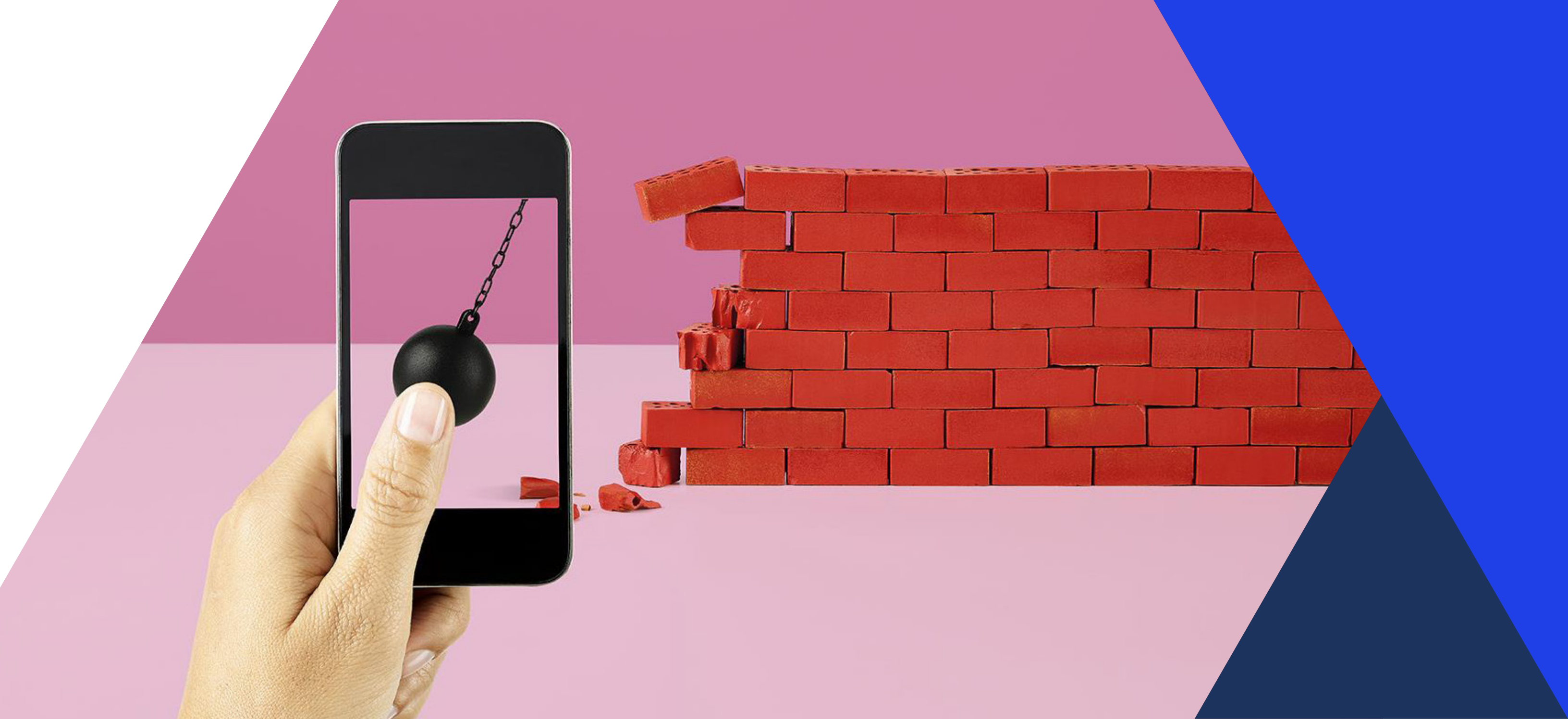
Every day we're working to help clients overcome these challenges and unlock the power of analytics. As part of that effort, we compiled this collection of articles to help you stay informed and stay ahead as analytics comes of age. We hope you enjoy it, learn from it, and share back with us your own perspectives on how these evolving technologies are changing our world.

Nicolaus Henke
Senior partner,
London

Noshir Kaka
Senior partner,
Mumbai

PART 1

Analytics grows up



The rise of ecosystems

Competing in a world of sectors without borders

Venkat Atluri, Miklós Dietz, and Nicolaus Henke

Digitization is causing a radical reordering of traditional industry boundaries. What will it take to play offense and defense in tomorrow's ecosystems?

Rakuten Ichiba is Japan's single largest online retail marketplace. It also provides loyalty points and e-money usable at hundreds of thousands of stores, virtual and real. It issues credit cards to tens of millions of members. It offers financial products and services that range from mortgages to securities brokerage. And the company runs one of Japan's largest online travel portals—plus an instant-messaging app, Viber, which has some 800 million users worldwide. Retailer? Financial company? Rakuten Ichiba is all that and more—just as Amazon and China's Tencent are tough to categorize as the former engages in e-commerce, cloud-computing, logistics, and consumer electronics, while the latter provides services ranging from social media to gaming to finance and beyond.

Organizations such as these—digital natives that are not defined or constrained by any one industry—may seem like outliers. How applicable to traditional industries is the notion of simultaneously competing in multiple sectors, let alone reimagining sector boundaries? We would be the first to acknowledge that opportunities to attack and to win across sectors vary considerably and that industry definitions have always been fluid: technological developments cause sectors to appear, disappear, and merge. Banking, for example, was born from the merger of money exchange, merchant banking, savings banking, and safety-deposit services, among others. Supermarkets unified previously separate retail subsectors into one big "grocery" category. Changes such as these created new competitors, shifted vast amounts of wealth, and reshaped significant parts of the economy. Before the term was in vogue, one could even say the shifts were "disruptive."

Yet there does appear to be something new happening here. The ongoing digital revolution, which has been reducing frictional, transactional costs for years, has accelerated recently with tremendous increases in electronic data, the ubiquity of mobile interfaces, and the growing power of artificial intelligence (AI). Together, these forces are reshaping customer expectations and creating the potential for virtually every sector with a distribution component to have its borders redrawn or redefined, at a more rapid pace than we have previously experienced.

Consider first how customer expectations are shifting. As Steve Jobs famously observed, "A lot of times, people don't know what they want until you show it to them." By creating a customer-centric, unified value proposition that extends beyond what end users could previously obtain (or, at least, could obtain almost instantly from one interface), digital pioneers are bridging the openings along the value chain, reducing customers' costs, providing them with new experiences, and whetting their appetites for more.

We've all experienced businesses that once seemed disconnected fitting together seamlessly and unleashing surprising synergies: look no farther than the phone in your pocket, your music and video in the cloud, the smart watch on your wrist, and the TV in your living room. Or consider the 89 million customers now accessing Ping An Good Doctor, where on a single platform run by the trusted Ping An insurance company they can connect with doctors not only for online bookings but to receive diagnoses and suggested treatments, often by exchanging pictures and videos. What used to take many weeks and multiple providers can now be done in minutes on one app.

Now nondigital natives are starting to think seriously about their cross-sector opportunities and existential threats that may lurk across boundaries. One example: We recently interviewed 300 CEOs worldwide, across 37 sectors, about advanced data analytics. Fully one-third of them had cross-sector dynamics at top of mind. Many worried, for instance, that "companies from other industries have clearer insight into my customers than I do." We've also seen conglomerates that until recently had thought of themselves as little more than holding companies taking the first steps to set up enterprise-wide consumer data lakes, integrate databases, and optimize the products, services, and insights they provide to their customers. Although these companies must of course abide by privacy laws—and even more, meet their users' expectations of trust—data sets and sources are becoming great unifiers and creating new, cross-sectoral competitive dynamics.

Do these dynamics portend a sea change for every company? Of course not. People will still stroll impromptu into neighborhood stores, heavy industry (with the benefit of technological advances, to be sure) will go on extracting and processing the materials essential to our daily lives, and countless other enterprises beyond the digital space will continue to channel the ingenuity of their founders and employees to serve a world of incredibly varied preferences and needs. It's obvious that digital will not—and cannot—change everything.

But it's just as apparent that its effects on the competitive landscape are already profound and that the stakes are getting higher. As boundaries between industry sectors continue to blur, CEOs—many of whose companies have long commanded large revenue pools within traditional industry lines—will face off against companies and industries they never previously viewed as competitors. This new environment will play out by new rules, require different capabilities, and rely to an extraordinary extent upon data. Defending your position will be mission critical, but so too will be attacking and capturing the opportunities across sectors before others get there first. To put it another way: within a decade, companies will define their business models not by how they play against traditional industry peers but by how effective they are in competing within rapidly emerging "ecosystems" comprising a variety of businesses from dimensionally different sectors.

A world of digital ecosystems

As the approaching contest plays out, we believe an increasing number of industries will converge under newer, broader, and more dynamic alignments: digital ecosystems. A world of ecosystems will be a highly customer-centric model, where users can enjoy an end-to-end experience for a wide range of products and services through a single access gateway, without leaving the ecosystem. Ecosystems will comprise diverse players who provide digitally accessed, multi-industry solutions. The relationships among these participants will be commercial and contractual, and the contracts (whether written, digital, or both) will formally regulate the payments or other considerations trading hands, the services provided, and the rules governing the provision of and access to ecosystem data.

Beyond just defining relationships among ecosystem participants, the digitization of many such arrangements is changing the boundaries of the company by reducing frictional costs associated with activities such as trading, measurement, and maintaining trust. More than 80 years ago, Nobel laureate Ronald Coase argued that companies establish their boundaries on the basis of transaction costs like these: when the cost of transacting for a product or service on the open market exceeds the cost of managing and coordinating the incremental activity needed to create that product or service internally, the company will perform the activity in-house. As digitization reduces transaction costs, it becomes economic for companies to contract out more activities, and a richer set of more specialized ecosystem relationships is facilitated.

Rising expectations

Those ecosystem relationships, in turn, are making it possible to better meet rising customer expectations. The mobile Internet, the data-crunching power of advanced analytics, and the maturation of artificial intelligence (AI) have led consumers to expect fully personalized solutions, delivered in milliseconds. Ecosystem orchestrators use data to connect the dots—by, for example, linking all possible producers with all possible customers, and, increasingly, by predicting the needs of customers before they are articulated. The more a company knows about its customers, the better able it is to offer a truly integrated, end-to-end digital experience and the more services in its ecosystem it can connect to those customers, learning ever more in the process. Amazon, among digital natives, and Centrica, the British utility whose Hive offering seeks to become a digital hub for controlling the home from any device, are early examples of how pivotal players can become embedded in the everyday life of customers.

For all of the speed with which sector boundaries will shift and even disappear, courting deep customer relationships is not a one-step dance. Becoming part of an individual's day-to-day experience takes time and, because digitization lowers switching costs and heightens price transparency, sustaining trust takes even longer. Over that time frame, significant surplus may shift to consumers—a phenomenon already underway, as digital players are destroying billions to create millions. It's also a process that will require deploying newer tools and technologies, such as using bots in multidevice environments and exploiting AI to build machine-to-machine capabilities. Paradoxically, sustaining customer relationships will depend as well on factors that defy analytical formulae: the power of a brand, the tone of one's message, and the emotions your products and services can inspire.

Strategic moves

The growing importance of customer-centricity and the appreciation that consumers will expect a more seamless user experience are reflected in the flurry of recent strategic moves of leading companies across the world. Witness Apple Pay; Tencent's and Alibaba's service expansions; Amazon's decisions to (among other things) launch Amazon Go, acquire Whole Foods, and provide online vehicle searches in Europe; and the wave of announcements from other digital leaders heralding service expansion across emerging ecosystems. Innovative financial players such as CBA (housing and B2B services), mBank (B2C marketplace), and Ping An (for health, housing, and autos) are mobilizing. So are telcos, including Telstra and Telus (each playing in the health ecosystem), and retailers such as Starbucks (with digital content, as well as seamless mobile payments and pre-ordering). Not to be left out are industrial companies such as GE (seeking to make analytics the new "core to the company") and Ford (which has started to redefine itself as "a mobility company and not just as a car and truck manufacturer").[1] We've also seen ecosystem-minded combinations such as Google's acquisition of Waze and Microsoft's purchase of LinkedIn. Many of these initiatives will seem like baby steps when we look back a decade from now, but they reveal the significance placed by corporate strategists on the emergence of a new world.

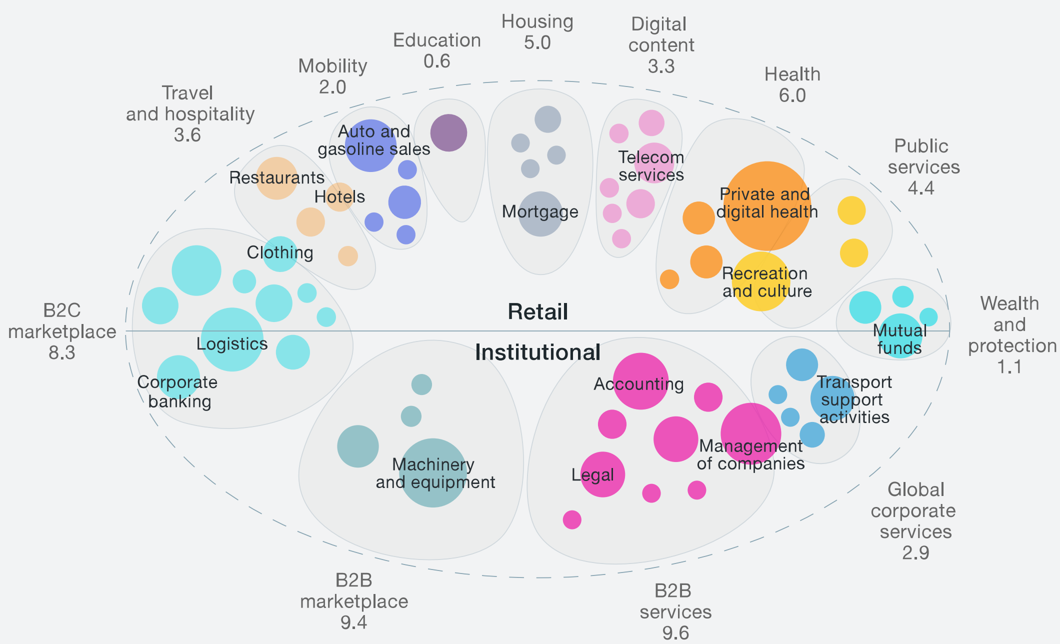
While it might be tempting to conclude as a governing principle that aggressively buying your way into new sectors is the secret spice for ecosystem success, massive combinations can also be recipes for massive value destruction. To keep your bearings in this new world, focus on what matters most—your core value propositions, your distinct competitive advantages, fundamental human and organizational needs, and the data and technologies available to tie them all together. That calls for thinking strategically about what you can provide your customers within a logically connected network of goods and services: critical building blocks of an ecosystem, as we've noted above.

1 See Nicolaus Henke, Ari Libarikian, and Bill Wiseman, "Straight talk about big data," *McKinsey Quarterly*, October 2016; and "Bill Ford charts a course for the future," *McKinsey Quarterly*, October 2014, both available on McKinsey.com.

Exhibit 1

New ecosystems are likely to emerge in place of many traditional industries by 2025.

Ecosystem illustration, estimated total sales in 2025,[1] $ trillion



[1]Circle sizes show approximate revenue pool sizes. Additional ecosystems are expected to emerge in addition to the those depicted; not all industries or subcategories are shown.

Source: IHS World Industry Service; Panorama by McKinsey; McKinsey analysis

Value at stake

Based on current trends, observable economic trajectories, and existing regulatory frameworks, we expect that within about a decade 12 large ecosystems will emerge in retail and institutional spaces. Their final shape is far from certain, but we suspect they could take something like the form presented in Exhibit 1.

The actual shape and composition of these ecosystems will vary by country and region, both because of the effects of regulations and as a result of more subtle, cultural customs and tastes. We already see in China, for example, how a large base of young tech-savvy consumers, a wide amalgam of low-efficiency traditional industries, and, not least, a powerful regulator have converged to give rise to leviathans such as Alibaba and Tencent—ideal for the Chinese market but not (at least, not yet) able to capture significant share in other geographies (see sidebar, "China by the numbers").

The value at stake is enormous. The World Bank projects the combined revenue of global businesses will be more than $190 trillion within a decade. If digital distribution (combining B2B and B2C commerce) represents about one-half of the nonproduction portion of the global economy by that time, the revenues that could, theoretically, be redistributed across traditional sectoral borders in 2025 would exceed $60 trillion—about 30 percent of world revenue pools that year. Under standard margin assumptions, this would translate to some $11 trillion in global profits, which, once we subtract approximately $10 trillion for cost of equity, amounts to $1 trillion in total economic profit.[2]

Snapshots of the future

Again, it's uncertain how much of this value will be reapportioned between businesses and consumers, let alone among industries, sectors, and individual companies, or whether and to what extent governments will take steps to weigh in. To a significant degree, many of the steps that companies are taking and contemplating are defensive in nature—fending off newer entrants, by using data and customer relationships to shore up their core. As incumbents and digital natives alike seek to secure their positions while building new ones, ecosystems are sure to evolve in ways that surprise us. Here is a quick look at developments underway in three of them.

Consumer marketplaces

By now, purchasing and selling on sites such as Alibaba, Amazon, and eBay is almost instinctive; retail has already been changed forever. But we expect that the very concept of a clearly demarcated retail sector will be radically altered within a decade. Three critical related factors are at work.

First, the frame of reference: what we think of now as one-off purchases will more properly be understood as part of a consumer's passage through time—the accumulation of purchases made from day to day, month to month, year to year, and ultimately the way those interact over a lifetime. Income and wealth certainly have predictive value for future purchases, but behavior matters even more. Choices to eat more healthily, for example, correlate to a likelihood for higher consumption of physical-fitness gear and services, and also to a more attractive profile for health and life insurers, which should result in more affordable coverage.

The second major factor, reinforcing the first, is the growing ability of data and analytics to transform disparate pieces of information about a consumer's immediate desires and behavior into insight about the consumer's broader needs. That requires a combination of capturing innumerable data points and turning them, within milliseconds, into predictive, actionable opportunities for both sellers and buyers. Advances in big data analytics, processing power, and AI are already making such connections possible.

This all generates a highly robust "network factor"—the third major force behind emerging consumer marketplaces. In a world of digital networks, consumer lenders, food and beverage providers, and telecom players will simultaneously coexist, actively partner, and aggressively move to capture share

2 Our conclusions, which we arrived at by analyzing 2025 profit pools from a number of different perspectives, are based upon several base expectations about the coming integrated network economy, including average profit margin and return on equity (for each, we used the world's top 800 businesses today, excluding manufacturing initiatives), as well as on the cost of equity (which we derived from more than 35,000 global companies based upon their costs of equity in January 2017).

China by the numbers

China has unique regulatory, demographic, and developmental features—particularly the simultaneity with which its economy has modernized and digitized—that are accelerating the blurring of sector borders. Still, the numbers speak for themselves and help suggest both the scale that digital ecosystems can quickly reach and the patterns likely to take hold elsewhere as ecosystem orchestrators in other countries stretch into roles approximating those played by Alibaba, Baidu, Ping An, and Tencent.

Alibaba

$120 billion
assets under management by Yu'E Bao[1]

175 million
total Alipay transactions in one day[2]

44%
of global mobile-wallet spending, achieved by Alipay[3]

Baidu

346 million
online users

130 million
users of Ping An Good Doctor[4]

25 million
unique visitors daily to autohome.com.cn

Tencent

889 million
WeChat users[5]

70 minutes
spent every day by average WeChat user[6]

61%
of users open WeChat more than ten times every day[7]

46 billion
"red packets" sent via WeChat for the lunar new year[8]

[1] As of September 2016.
[2] As of August 2016.
[3] In 2016; see *Global Payments Report 2016*, Worldpay, November 8, 2016, worldpay.com.
[4] As of March 2017.
[5] As of Q4 2016.
[6] As of March 2016.
[7] As of June 2016.
[8] For Lunar New Year falling in 2017; see "WeChat users send 46 billion digital red packets over Lunar New Year—Xinhua," Reuters, February 6, 2017, reuters.com.

Large Chinese players have expanded their digital presence by ‘land grabbing.’

Selected examples

Alibaba Baidu Tencent

2000

Market, consumption

Search

Messaging

2017

Market, consumption

Alibaba.com, Taobao, Tmall

Baidu Wei Gou, Wanda e-commerce

We Store, Xi Yuan

Search

Baidu Map, Baidu Search

Sogou

Messaging

QQ, WeChat

Dining

Ele.me

Baidu Nuomi, Baidu Takeout Delivery

Meituan-Dianping

Entertainment, gaming

Alibaba Games, Alibaba Music, Alibaba Picture

Baidu Games, Baidu Music, Baidu Video, iQIYI

QQ Music, Tencent Games, Tencent Video

Finance

Ant Financial Services Group

Baidu Consumer Credit, Baidu Wallet, Baidu Wealth Management

Caifutong, Tenpay, WeBank

Healthcare

Alihealth

Ding Xiang Yuan

News, encyclopedia

Baidu Baike, Baidu News

Transportation

Didi Chuxing[1]

[1]Formed by merger of Didi Dache (backed by Tencent) and Kuaidi Dache (backed by Alibaba) and acquisition of Uber (backed by Baidu).

Source: Company websites

from one another. And while digitization may offer the sizzle, traditional industries still have their share of the steak. These businesses not only provide the core goods and services that end users demand, but often also have developed relationships with other businesses along the value chain and, most important, with the end users themselves. Succeeding in digital marketplaces will require these companies to stretch beyond their core capabilities, to be sure, but if they understand the essentials of what's happening and take the right steps to secure and expand their relationships, nondigital businesses can still hold high ground when the waves of change arrive.

B2B services

The administrative burdens of medium, small, and microsize companies are both cumbersome and costly. In addition to managing their own products and services, these businesses (like their larger peers) must navigate a slew of necessary functions, including human resources, tax planning, legal services, accounting, finance, and insurance.

Today, each of these fields exists as an independent sector, but it's easy to imagine them converging within a decade on shared, cloud-based platforms that will serve as one-stop shops. With so many service providers available at the ease of a click, all with greater transparency on price, performance, and reputation, competition will ramp up, and established players can anticipate more challengers from different directions. At the same time, it's likely that something approaching a genuine community will develop, with businesses being able to create partnerships and tap far more sophisticated services than they can at present—including cash-planning tools, instant credit lines, and tailored insurance.

Already, we can glimpse such innovations starting to flourish in a range of creative solutions. Idea Bank in Poland, for example, offers "idea hubs" and applications such as e-invoicing and online factoring. ING's commercial platform stretches beyond traditional banking services to include (among other things) a digital loyalty program and crowdfunding. And Lloyds Bank's Business Toolbox includes legal assistance, online backup, and email hosting. As other businesses join in, we expect the scope and utility of this space to grow dramatically.

Mobility

Finally, consider personal mobility, which encompasses vehicle purchase and maintenance management, ridesharing, carpooling, traffic management, vehicle connectivity, and much more. The individual pieces of the mobility puzzle are starting to become familiar, but it's their cumulative impact that truly shows the degree to which industry borders are blurring (Exhibit 2).

Emerging priorities for the borderless economy

These glimpses of the future are rooted in the here and now, and they are emblematic of shifts underway in most sectors of the economy—including, more likely than not, yours. We hope this article is a useful starting point for identifying potential industry shifts that could be coming your way. Recognition is the first step, and then you need a game plan for a world of sectors without borders. The following four priorities are critical:

- **Adopt an ecosystem mind-set.** The landscape described in this article differs significantly from the one that still dominates most companies' business planning and operating

Exhibit 2

Different sectors come into play at every stage of the mobility ecosystem.

Source: Panorama by McKinsey

approaches. Job one for many companies is to broaden their view of competitors and opportunities so that it is truly multisectoral, defines the ecosystems and industries where change will be fastest, and identifies the critical new sources of value most meaningful for an expanding consumer base. In essence, you must refine your "self vision" by asking yourself and your top team questions such as: "What surprising, disruptive boundary shifts can we imagine—and try to get ahead of?" and "How can we turn our physical assets and long-established customer relationships into genuine consumer insights to secure what we have and stake out an advantage over our competitors—including the digital giants?" That shift will necessarily involve an important organizational component, and leaders should expect some measure of internal resistance, particularly when existing business goals, incentives, and performance-management principles do not accord with new strategic priorities. It will also, of course, require competitive targeting beyond the four walls of your company. But resist the impulse to just open up your acquisition checkbook. The combinations that make good sense will be part of a rational answer to perennial strategic questions about where and how your company needs to compete—playing out on an expanding field.

- **Follow the data.** In our borderless world, data are the coins of the realm. Competing effectively means both collecting large amounts of data and developing capabilities for storing, processing, and translating the data into actionable business insights. A critical goal for most companies is data diversity—achieved, in part, through partnerships—which will enable you to pursue ever-finer microsegmentation and create more value in more ecosystems. Information from telecommunications-services players, for example, can help banks to engage their customers and make a variety of commercial decisions more effectively. Deeper data insights are finally beginning to take ideas that had always seemed good but too often fell short of their potential to turn into winning models. Consider loyalty cards: by understanding customers better, card providers such as Nectar, the largest loyalty program in the United Kingdom, and Plenti, a rewards programs introduced by American Express, can connect hundreds of companies of all sizes and across multiple industries to provide significant savings for consumers and new growth opportunities for the businesses that serve them. Meanwhile, the cost of sharing data is falling as cloud-based data stores proliferate and AI makes it easier to link data sets to individual customers or segments. Better data can also support analytically driven scenario planning to inform how ecosystems will evolve, at which points along the value chain your data can create value, and whether or where you can identify potential "Holy Grail" data assets. What data points and sources are critical to your business? How many do you have? What can you do to acquire or gain access to the rest? You should be asking your organization questions like these regularly.

- **Build emotional ties to customers.** If blurring sector boundaries are turning data into currency, customer ownership is becoming the ultimate prize. Companies that lack strong customer connections run the risk of disintermediation and perhaps of becoming "white-label back offices" (or production centers), with limited headroom to create or retain economic surplus. Data (to customize offerings), content (to capture the attention of customers), and digital engagement models (to create seamless customer journeys that solve customer pain points) can all help you build emotional connections with customers and occupy attractive roles in critical ecosystems. You should continually be asking your organization, "What's our plan for using data, content, and digital-engagement tools to connect emotionally with customers?" and "What else can we provide, with simplicity and speed, to strengthen our consumer bond?" After all, Google's launch of initiatives such as Chrome and Gmail, and Alibaba's introduction of enterprises such as Alipay and the financial platform Yu'E Bao, weren't executed merely because they already had a huge customer base and wanted to capture new sources of revenue (although they did succeed in doing so). They took action to help ensure they would keep—and expand—that huge customer base.

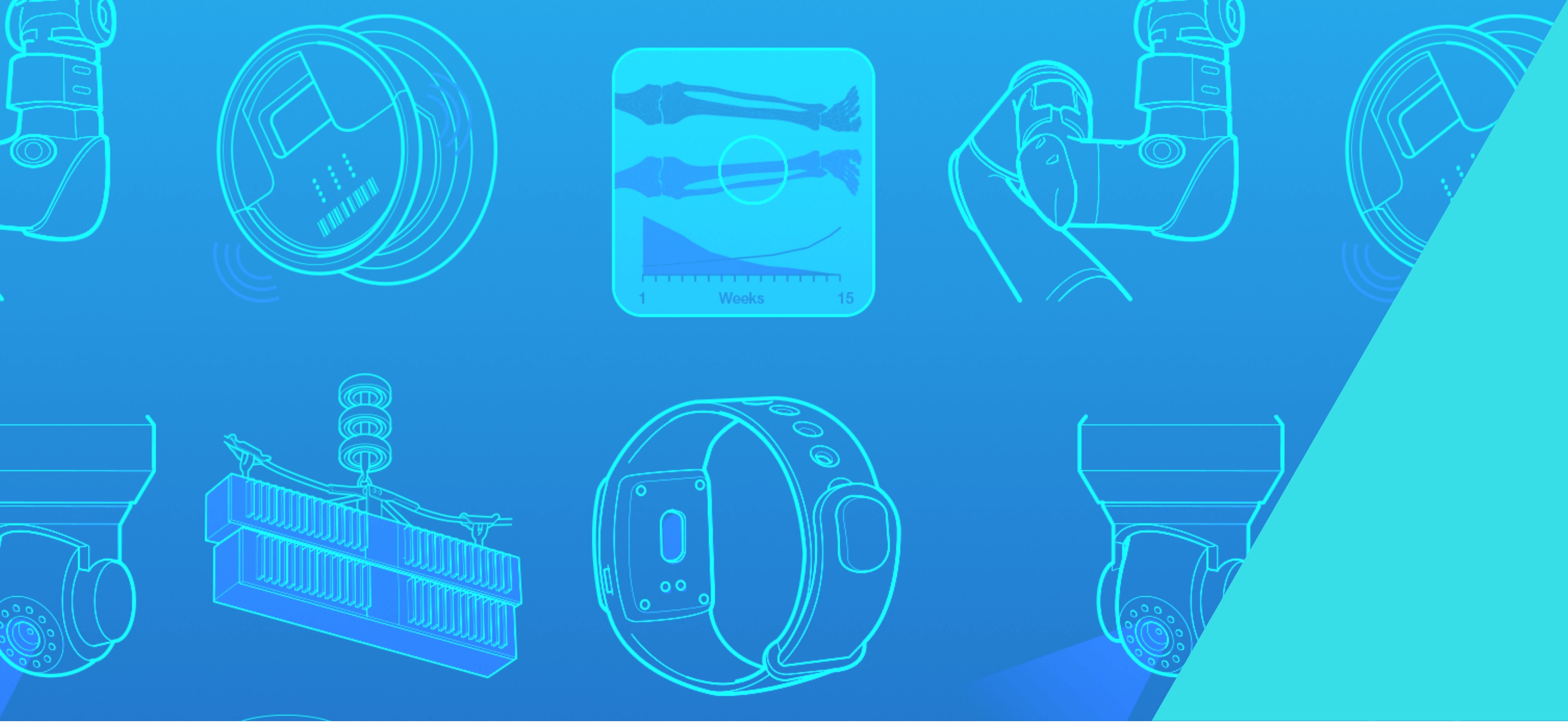
- **Change your partnership paradigm.** Given the opportunities for specialization created by an ecosystem economy, companies need more and different kinds of partners. In at least a dozen markets worldwide—including Brazil, Turkey, and several countries in Asia, where in many respects data are currently less robust than they are in other regions—we're seeing a new wave of partnership energy aimed at making the whole greater than the sum of its parts. Regardless of your base geography, core industry, and state of data readiness, start by asking what white spaces you need to fill, what partners can best help with those gaps, and what "gives" and "gets" might be mutually beneficial. You'll also need to think about how to create an infrastructural and operational framework that invites a steady exchange with outside entities of data, ideas, and services to fuel innovation. Don't forget about the implications for your information architecture, including the application programming interfaces (APIs) that will enable critical external linkages, and don't neglect the possibility that you may need to enlist a more natural integrator from across your partnerships, which could include a company more appropriate for the role, such as a telco, or a third-party provider that can more effectively connect nondigital natives. And don't assume that if you were to acquire a potential partner, you'd necessarily be adding and sustaining their revenues on a dollar-for-dollar basis over the long term.

▼▼▼

No one can precisely peg the future. But when we study the details already available to us and think more expansively about how fundamental human needs and powerful technologies are likely to converge going forward, it is difficult to conclude that tomorrow's industries and sector borders will look like today's. Massive, multi-industry ecosystems are on the rise, and enormous amounts of value will be on the move. Companies that have long operated with relative insularity in traditional industries may be most open to cross-boundary attack. Yet with the right strategy and approach, leaders can exploit new openings to go on offense, as well. Now is the time to take stock and to start shaping nascent opportunities.

Venkat Atluri is a senior partner in McKinsey's Chicago office, **Miklós Dietz** is a senior partner in the Vancouver office, and **Nicolaus Henke** is a senior partner in the London office.

The authors wish to thank Miklos Radnai, Global Head of McKinsey's Ecosystems Working Group, and McKinsey's Tamas Kabay, Somesh Khanna, and Istvan Rab for their contributions to this article.



The arrival of artificial intelligence

Artificial intelligence is getting ready for business, but are businesses ready for AI?

Tera Allas, Jacques Bughin, Michael Chui, Peter Dahlström, Eric Hazan, Nicolaus Henke, Sree Ramaswamy, and Monica Trench

Companies new to the space can learn a great deal from early adopters who have invested billions in AI and are now beginning to reap a range of benefits.

Claims about the promise and peril of artificial intelligence are abundant, and growing. AI, which enables machines to exhibit humanlike cognition, can drive our cars or steal our privacy, stoke corporate productivity or empower corporate spies. It can relieve workers of repetitive or dangerous tasks or strip them of their livelihoods. Twice as many articles mentioned AI in 2016 as in 2015, and nearly four times as many as in 2014.[1] Expectations are high.

AI has been here before. Its history abounds with booms and busts, extravagant promises, and frustrating disappointments. Is it different this time? New analysis suggests yes: AI is finally

1 Factiva.

starting to deliver real-life business benefits. The ingredients for a breakthrough are in place. Computer power is growing significantly, algorithms are becoming more sophisticated, and, perhaps most important of all, the world is generating vast quantities of the fuel that powers AI—data. Billions of gigabytes of it every day.

Companies at the digital frontier—online firms and digital natives such as Google and Baidu—are betting vast amounts of money on AI. We estimate between $20 billion and $30 billion in 2016, including significant M&A activity. Private investors are jumping in, too. We estimate that venture capitalists invested $4 billion to $5 billion in AI in 2016, and private equity firms invested $1 billion to $3 billion. That is more than three times as much as in 2013. An additional $1 billion of investment came from grants and seed funding.

For now, though, most of the news is coming from the suppliers of AI technologies. And many new uses are only in the experimental phase. Few products are on the market or are likely to arrive there soon to drive immediate and widespread adoption. As a result, analysts remain divided as to the potential of AI: some have formed a rosy consensus about AI's potential while others remain cautious about its true economic benefit. This lack of agreement is visible in the large variance of current market forecasts, which range from $644 million to $126 billion by 2025.[2] Given the size of investment being poured into AI, the low estimate would indicate that we are witnessing another phase in a boom-and-bust cycle.

Our business experience with AI suggests that this bust scenario is unlikely. In order to provide a more informed view, we decided to perform our own research into how users are adopting AI technologies. Our research offers a snapshot of the current state of the rapidly changing AI industry. To begin, we examine the investment landscape, including firms' internal investment in R&D and deployment, large corporate M&A, and funding from venture capital (VC) and private equity (PE) firms. We then combine use-case analyses and our AI adoption and use survey of C-level executives at more than 3,000 companies to understand how companies use AI technologies today.

AI generally refers to the ability of machines to exhibit humanlike intelligence—for example, solving a problem without the use of hand-coded software containing detailed instructions. There are several ways to categorize AI technologies, but it is difficult to draft a list that is mutually exclusive and collectively exhaustive, because people often mix and match several technologies to create solutions for individual problems. These creations sometimes are treated as independent technologies, sometimes as subgroups of other technologies, and sometimes as applications. Some frameworks group AI technologies by basic functionality, such as text, speech, or image recognition, and some group them by business applications such as commerce or cybersecurity.[3]

---

2 Tractica; Transparency Market Research.

3 Gil Press, "Top 10 hot artificial intelligence (AI) technologies," Forbes.com, January 23, 2017; "AI100: The artificial intelligence start-ups redefining industries," CBinsights.com, January 11, 2017.

Trying to pin down the term more precisely is fraught for several reasons: AI covers a broad range of technologies and applications, some of which are merely extensions of earlier techniques and others that are wholly new. Also, there is no generally accepted theory of "intelligence," and the definition of machine "intelligence" changes as people become accustomed to previous advances.[4] Tesler's theorem, attributed to the computer scientist Larry Tesler, asserts that "AI is whatever hasn't been done yet."[5]

The AI technologies we consider in this paper are what is called "narrow" AI, which performs one narrow task, as opposed to artificial general intelligence, or AGI, which seeks to be able to perform any intellectual task that a human can do. We focus on narrow AI because it has near-term business potential, while AGI has yet to arrive.[6]

In this report, we focus on the set of AI technology systems that solve business problems. We have categorized these into five technology systems that are key areas of AI development: robotics and autonomous vehicles, computer vision, language, virtual agents, and machine learning, which is based on algorithms that learn from data without relying on rules-based programming in order to draw conclusions or direct an action. Some are related to processing information from the external world, such as computer vision and language (including natural language processing, text analytics, speech recognition, and semantics technology); some are about learning from information, such as machine learning; and others are related to acting on information, such as robotics, autonomous vehicles, and virtual agents, which are computer programs that can converse with humans. Machine learning and a subfield called deep learning are at the heart of many recent advances in artificial intelligence applications and have attracted a lot of attention and a significant share of the financing that has been pouring into the AI universe—almost 60 percent of all investment from outside the industry in 2016.

Artificial intelligence's roller-coaster ride to today

Artificial intelligence, as an idea, first appeared soon after humans developed the electronic digital computing that makes it possible. And, like digital technology, artificial intelligence, or AI, has ridden waves of hype and gloom—with one exception: AI has not yet experienced wide-scale commercial deployment (see sidebar, "Fits and starts: A history of artificial intelligence").

That may be changing. Machines powered by AI can today perform many tasks—such as recognizing complex patterns, synthesizing information, drawing conclusions, and forecasting—that not long ago were assumed to require human cognition. And as AI's capabilities have dramatically expanded, so has its utility in a growing number of fields. At the same time, it is worth remembering that machine learning has limitations. For example, because the systems are

4 Marvin Minsky, "Steps toward artificial intelligence," *Proceedings of the IRE*, volume 49, number 1, January 1961; Edward A. Feigenbaum, *The art of artificial intelligence: Themes and case studies of knowledge engineering*, Stanford University Computer Science Department report number STAN-CS-77–621, August 1977; Allen Newell, "Intellectual issues in the history of artificial intelligence," in *The Study of Information: Interdisciplinary messages*, Fritz Machlup and Una Mansfield, eds., John Wiley and Sons, 1983.

5 Douglas R. Hofstadter, *Gödel, Escher, Bach: An eternal golden braid*, Basic Books, 1979. Hofstadter writes that he gave the theorem its name after Tesler expressed the idea to him firsthand. However, Tesler writes in his online CV that he actually said, "Intelligence is whatever machines haven't done yet."

6 William Vorhies, "Artificial general intelligence—the Holy Grail of AI," DataScienceCentral.com, February 23, 2016.

trained on specific data sets, they can be susceptible to bias; to avoid this, users must be sure to train them with comprehensive data sets. Nevertheless, we are seeing significant progress.

Fits and starts: A history of artificial intelligence

The idea of computer-based artificial intelligence dates to 1950, when Alan Turing proposed what has come to be called the Turing test: Can a computer communicate well enough to persuade a human that it, too, is human?[1] A few months later, Princeton students built the first artificial neural network, using 300 vacuum tubes and a war-surplus gyropilot.[2]

The term "artificial intelligence" was coined in 1955, to describe the first academic conference on the subject, at Dartmouth College. That same year, researchers at the Carnegie Institute of Technology (now Carnegie Mellon University) produced the first AI program, Logic Theorist.[3] Advances followed often through the 1950s: Marvin Lee Minsky founded the Artificial Intelligence Laboratory at MIT, while others worked on semantic networks for machine translation at Cambridge and self-learning software at IBM.[4]

Funding slumped in the 1970s as research backers, primarily the US government, tired of waiting for practical AI applications and cut appropriations for further work.[5] The field was fallow for the better part of a decade.

University researchers' development of "expert systems"—software programs that assess a set of facts using a database of expert knowledge and then offer solutions to problems—revived AI in the 1980s.[6] Around this time, the first computer-controlled autonomous vehicles began to appear.[7] But this burst of interest preceded another AI "winter."

Interest in AI boomed again in the 21st century as advances in fields such as deep learning, underpinned by faster computers and more data, convinced investors and researchers that it was practical—and profitable—to put AI to work.[8]

1 A. M. Turing, "Computing machinery and intelligence," *Mind*, volume 49, number 236, October 1950.

2 Jeremy Bernstein, "A.I.," *The New Yorker*, December 14, 1981.

3 Leo Gugerty, "Newell and Simon's Logic Theorist: Historical background and impact on cognitive modeling," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, volume 50, issue 9, October 2006.*

4 "The IBM 700 Series: Computing comes to business," IBM Icons of Progress, March 24, 2011.

5 Michael Negnevitsky, *Artificial intelligence: A guide to intelligent systems*, Addison-Wesley, 2002.

6 Edward A. Feigenbaum, "Expert systems in the 1980s," working paper, 1980.

7 Hans P. Moravec, "The Stanford Cart and the CMU Rover," *Proceedings of the IEEE*, volume 71, issue 7, July 1983; Tom Vanderbilt, "Autonomous cars through the ages," Wired.com, February 6, 2012.

8 Bruce G. Buchanan, "A (very) brief history of artificial intelligence," *AI Magazine*, volume 26, number 4, Winter 2005.

These advances have allowed machine learning to be scaled up since 2000 and used to drive deep learning algorithms, among other things. The advances have been facilitated by the availability of large and diverse data sets, improved algorithms that find patterns in mountains of data, increased R&D financing, and powerful graphics processing units (GPUs), which have brought new levels of mathematical computing power. GPUs, which are specialized integrated circuits originally developed for video games, can process images 40 to 80 times faster than the fastest versions available in 2013. Advances in the speed of GPUs have enabled the training speed of deep learning systems to improve five- or sixfold in each of the last two years. More data—the world creates about 2.2 exabytes, or 2.2 billion gigabytes, of it every day—translates into more insights and higher accuracy because it exposes algorithms to more examples they can use to identify correct and reject incorrect answers. Machine learning systems enabled by these torrents of data have reduced computer error rates in some applications—for example, in image identification—to about the same as the rate for humans.

AI investment is growing rapidly, but commercial adoption is lagging

Tech giants and digital native companies such as Amazon, Apple, Baidu, and Google are investing billions of dollars in the various technologies known collectively as artificial intelligence. They see that the inputs needed to enable AI to finally live up to expectations—powerful computer hardware, increasingly sophisticated algorithmic models, and a vast and fast-growing inventory of data—are in place. Indeed, internal investment by large corporations dominates: we estimate that this amounted to $18 billion to $27 billion in 2016; external investment (from VCs, PE firms, M&A, grants, and seed funding) was around $8 billion to $12 billion (Exhibit 1).[7]

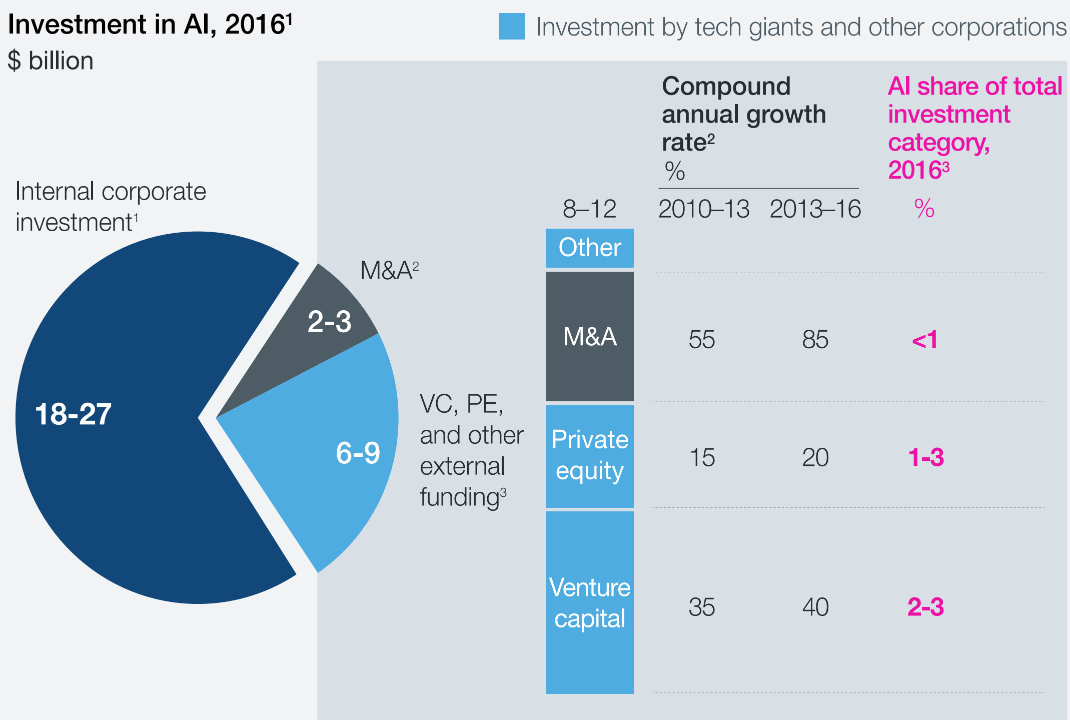
But for all the recent investment, the scope of AI deployment has been limited so far. That is partly due to the fact that one beneficiary of that investment, internal R&D, is largely focused on improving the firms' own performance. But it is also true that there is only tepid demand for artificial intelligence applications for businesses, partly due to the relatively slow pace of digital and analytics transformation of the economy. Our survey of more than 3,000 businesses around the world found that many business leaders are uncertain about what exactly AI can do for them, where to obtain AI-powered applications, how to integrate them into their companies, and how to assess the return on an investment in the technology.

Most of the investment in AI has consisted of internal spending—R&D and deployment—by large, cash-rich digital native companies. What is the large corporate investment in AI focused on? Bigger companies, such as Apple, Baidu, and Google, are working on suites of technologies internally, but vary in the breadth and focus of their AI investment. Amazon is working on robotics and speech recognition, Salesforce on virtual agents and machine learning. BMW, Tesla, and Toyota are among the manufacturers making sizable commitments in robotics and machine

7 Internal investment includes research and development, talent acquisition, cooperation with scientific institutions, and joint ventures with other companies done by corporations. External investment includes mergers and acquisitions, private equity funding, venture capital financing, and seed funds and other early-stage investing. The estimates of external investment are based on data available in the Capital IQ, PitchBook, and Dealogic databases. Provided values are estimates of annual investment in AI, assuming that all registered deals were completed within the year of transaction. Internal investment is estimated based on the ratio of AI spend to revenue for the top 35 high-tech and advanced manufacturing companies focused on AI technologies.

Exhibit 1

Technology giants dominate investment in AI



1 Estimate of 2016 spend by corporations to develop and deploy AI-based products. Calculated for top 35 high-tech and advanced manufacturing companies investing in AI. Estimate is based on the ratio of AI spend to total revenue calculated for a subset of the 35 companies.
2 VC value is an estimate of VC investment in companies primarily focused on AI. PE value is an estimate of PE investment in AI-related companies. M&A value is an estimate of AI deals done by corporations. “Other” refers to grants and seed fund investments. Includes only disclosed data available in databases, and assumes that all registered deals were completed within the year of transaction. Compound annual growth rate values rounded.
3 M&A and PE deals expressed by volume; VC deals expressed by value.

Source: Capital IQ; Pitchbook; Dealogic; S&P; McKinsey Global Institute analysis

learning for use in driverless cars. Toyota, for example, set aside $1 billion to establish a new research institute devoted to AI for robotics and driverless vehicles.[8] Industrial giants such as ABB, Bosch, GE, and Siemens also are investing internally, often in machine learning and robotics, seeking to develop specific technologies related to their core businesses. IBM has pledged to invest $3 billion to make its Watson cognitive computing service a force in the Internet

8 Craig Trudell and Yuki Hagiwara, “Toyota starts $1 billion center to develop cars that don’t crash,” Bloomberg.com, November 6, 2015.

of Things.[9] Baidu has invested $1.5 billion in AI research over the last 2½ years. This is in addition to $200 million it committed to a new in-house venture capital fund, Baidu Venture.[10]

At the same time, big tech companies have been actively buying AI start-ups, not just to acquire technology or clients but to secure qualified talent. The pool of true experts in the field is small, and Alibaba, Amazon, Facebook, Google, and other tech giants have hired many of them. Companies have adopted M&A as a way to sign up top talent, a practice known as "acqui-hiring," for sums that typically work out to $5 million to $10 million per person. The shortage of talent and cost of acquiring it are underlined by a recent report that companies are seeking to fill 10,000 AI-related jobs and have budgeted more than $650 million for salaries.[11]

Overall, corporate M&A is the fastest-growing external source of funding for AI companies, increasing in terms of value at a compound annual growth rate of over 80 percent from 2013 to 2016, based on our estimates. Leading high-tech companies and advanced manufacturers have closed more than 100 M&A deals since 2010. Google completed 24 transactions in that time, including eight in computer vision and seven in language processing. Apple, the second-most-active acquirer, has closed nine, split evenly among computer vision, machine learning, and language processing.

Companies are also expanding their search for talent abroad. Facebook, for instance, is opening an AI lab in Paris that will supplement similar facilities in New York and Silicon Valley—and make it easier for the company to recruit top researchers in Europe.[12] Google recently invested $4.5 million in the Montreal Institute for Learning Algorithms, a research lab at the University of Montreal; Intel donated $1.5 million to establish a machine learning and cybersecurity research center at Georgia Tech; and NVIDIA is working with the National Taiwan University to establish an AI laboratory in Taipei.[13]

The buzz over AI has grown loud enough to encourage venture capital and private equity firms to step up their investment in AI. Other external investors, such as angel funds and seed incubators, also are active. We estimate total annual external investment was $8 billion to $12 billion in 2016.[14]

---

9 "IBM invests to lead global Internet of Things market—shows accelerated client adoption," IBM press release, October 3, 2006.

10 Phoenix Kwong, "Baidu launches $200m venture capital unit focused on artificial intelligence," *South China Morning Post*, September 13, 2016.

11 "U.S. companies raising $1 billion or more to fuel artificial intelligence (AI) development: Looking to staff 10,000+ openings, cites new Paysa research," Paysa press release, April 18, 2017.

12 Cade Metz, "Facebook opens a Paris lab as AI research goes global," Wired.com, June 2, 2015.

13 Cade Metz, "Google opens Montreal AI lab to snag scarce global talent," Wired.com, November 12, 2015; "Georgia Tech launches new research on the security of machine-learning systems," Georgia Institute of Technology press release, October 31, 2016; "NVIDIA collaborates with Taipei Tech to establish Embedded GPU Joint Lab," National Taipei University of Technology press release, September 4, 2014.

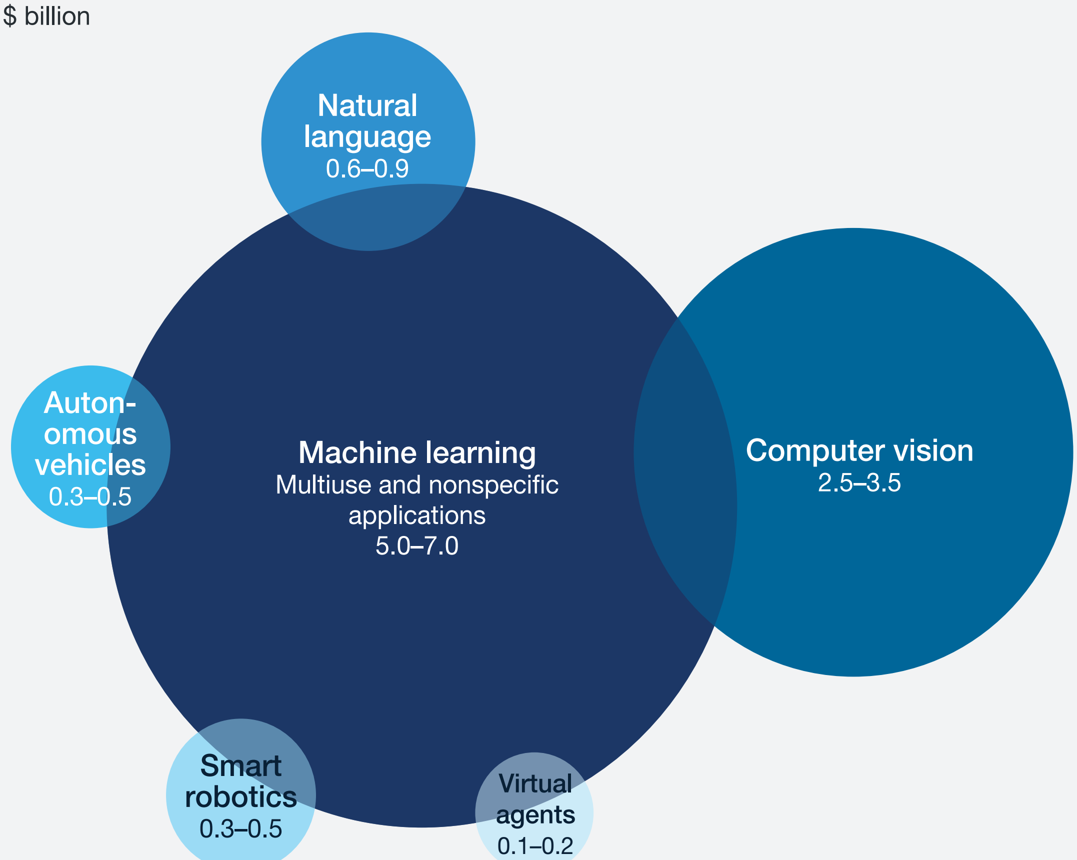
14 Estimates of external investment in AI vary widely because measurement standards vary. For example, Venture Scanner puts total funding of AI-related start-ups in 2016 at $2.5 billion, while Goldman Sachs estimates that the venture capital sector alone made $13.7 billion of AI-related investment that year.

Machine learning attracted almost 60 percent of that investment, most likely because it is an enabler for so many other technologies and applications, such as robotics and speech recognition (Exhibit 2). In addition, investors are drawn to machine learning because, as has long been the case, it is quicker and easier to install new code than to rebuild a robot or other machine that runs the software. Corporate M&A in this area is also growing fast, with a compound annual growth rate of around 80 percent from 2013 through 2016.

Exhibit 2

Machine learning received the most investment, although boundaries between technologies are not clear-cut

External investment in AI-focused companies by technology category, 2016[1]
$ billion



1 Estimates consist of annual VC investment in AI-focused companies, PE investment in AI-related companies, and M&A by corporations. Includes only disclosed data available in databases, and assumes that all registered deals were completed within the year of transaction.

Source: Capital IQ; Pitchbook; Dealogic; McKinsey Global Institute analysis

Investment in AI is still in the early stages and relatively small compared with the investment in the digital revolution. Artificial intelligence, for example, attracted 2 to 3 percent of all VC funding by value in 2016, while information technology in general soaked up 60 percent. AI also was a small fraction—1 to 3 percent—of all investment by PE firms in 2016.[15] But AI investment is growing fast.

From 2013 through 2016, external investment in AI technologies had a compound annual growth rate of almost 40 percent. That compares with 30 percent from 2010 through 2013. Not only are deals getting bigger and more numerous, but they require fewer participants to complete the financing. This suggests that investors are growing more confident in the sector and may have a better understanding of the technology and its potential.

However, for the most part, investors are still waiting for their investments to pay off. Only 10 percent of start-up companies that consider machine learning to be a core business say they generate revenue, according to PitchBook. Of those, only half report more than $50 million in revenue. Moreover, external investment remains highly concentrated geographically, dominated by a few technology hubs in the United States and China, with Europe lagging far behind.

Firms and industries already on the digital frontier are adopting AI, but others are hesitant to act

Investors are pouring billions of dollars into AI companies based on the hope that a market of AI adopters will develop fairly quickly and will be willing to pay for AI infrastructure, platforms, and services. Clearly, Amazon, Google, and other digital natives are investing for their own applications, such as optimizing searches and personalizing marketing. But getting a sense of how much traditional companies in healthcare, retail, and telecom are spending on AI is not easy. For this reason, we conducted a survey to understand this situation in more depth.

In general, few companies have incorporated AI into their value chains at scale; a majority of companies that had some awareness of AI technologies are still in experimental or pilot phases. In fact, out of the 3,073 respondents, only 20 percent said they had adopted one or more AI-related technology at scale or in a core part of their business.[16] Ten percent reported adopting more than two technologies, and only 9 percent reported adopting machine learning.[17]

Even this may overstate the commercial demand for AI at this point. Our review of more than 160 global use cases across a variety of industries found that only 12 percent had progressed beyond

15 It is worth noting that VC funds were focusing on AI technology when choosing investments, while PE funds were investing in AI-related companies.

16 Survey results throughout this discussion paper are weighted for firm size; “20 percent of firms” indicates firms representing 20 percent of the workforce.

17 The eight technologies are natural-language processing, natural-language generation, speech recognition, machine learning, decision management, virtual agents, robotics process automation, and computer vision. The five technology systems are robotics and autonomous vehicles, computer vision, language, virtual agents, and machine learning.

the experimental stage. Commercial considerations can explain why some companies may be reluctant to act. In our survey, poor or uncertain returns were the primary reason for not adopting reported by firms, especially smaller firms. Regulatory concerns have also become much more important.

As with every new wave of technology, we expect to see a pattern of early and late adopters among sectors and firms. We uncover six features of the early pattern of AI adoption, which is broadly in line with the ways companies have been adopting and using the recent cohort of digital technologies. Not coincidentally, the same players who were leaders in that earlier wave of digitization are leading in AI—the next wave.

The first feature is that early AI adopters are from sectors already investing at scale in related technologies, such as cloud services and big data. Those sectors are also at the frontier of digital assets and usage.[18] This is a crucial finding, as it suggests that there is limited evidence of sectors and firms catching up when it comes to digitization, as each new generation of tech builds on the previous one.

Second, independently of sectors, large companies tend to invest in AI faster at scale. This again is typical of digital adoption, in which, for instance, small and midsized businesses have typically lagged behind in their decision to invest in new technologies.

Third, early adopters are not specializing in one type of technology. They go broader as they adopt multiple AI tools addressing a number of different use cases at the same time.

Fourth, companies investing at scale do it close to their core business.

Fifth, early adopters that adopt at scale tend to be motivated as much by the upside growth potential of AI as they are by cutting costs. AI is not only about process automation, but is also used by companies as part of major product and service innovation. This has been the case for early adopters of digital technologies and suggests that AI-driven innovation will be a new source of productivity and may further expand the growing productivity and income gap between high-performing firms and those left behind.[19]

Finally, strong executive leadership goes hand in hand with stronger AI adoption. Respondents from firms that have successfully deployed an AI technology at scale tended to rate C-suite support nearly twice as high as those from companies that had not adopted any AI technology.

18 *Digital Europe: Pushing the frontier, capturing the benefits*, McKinsey Global Institute, June 2016; *Digital America: A tale of the haves and have-mores*, McKinsey Global Institute, December 2015.

19 Rosina Moreno and Jordi Suriñach, “Innovation adoption and productivity growth: Evidence for Europe,” working paper, 2014; Jacques Bughin and Nicolas van Zeebroeck, “The right response to digital disruption,” *MIT Sloan Management Review*, April 2017.

Early-adopting sectors are closer to the digital frontier

Sector-by-sector adoption of AI is highly uneven right now, reflecting many features of digital adoption more broadly. Our survey found that larger companies and industries that adopted digital technologies in the past are more likely to adopt AI. For them, AI is the next wave of digitization.

This pattern in the adoption of technology is not new—we saw similar behavior in firms adopting enterprise social technologies.[20] But this implies that, at least in the near future, AI deployment is likely to accelerate at the digital frontier, expanding the gap between adopters and laggards across companies, industries, and geographic regions.

The leading sectors include some that MGI's Industry Digitization Index found at the digital frontier, namely high tech and telecom and financial services.[21] These are industries with long histories of digital investment. They have been leaders in developing or adopting digital tools, both for their core product offerings and for optimizing their operations. However, even these sectors are far behind in AI adoption when compared with overall digitization (Exhibit 3).

Automotive and assembly is also highly ranked. It was one of the first sectors that implemented advanced robotics at scale for manufacturing, and today is also using AI technologies to develop self-driving cars.

In the middle are less digitized industries, including resources and utilities, personal and professional services, and building materials and construction. A combination of factors may account for this. These sectors have been slow to employ digital tools generally, except for some parts of the professional services industry and large construction companies. They are also industries in which innovation and productivity growth has lagged, potentially in part due to their domestic focus. Some of these sectors have a particularly high number of small firms—an important predictor for AI adoption, as explored below.

Toward the bottom of the pack for now are traditionally less digital fields such as education and healthcare. Despite ample publicity about cutting-edge AI applications in these industries, the reality is that uptake appears to be low so far. Weaker adoption reflects the particular challenges faced in these sectors. In healthcare, for example, practitioners and administrators acknowledge the potential for AI to reduce costs but quickly add that they believe that regulatory concerns and customer acceptance will inhibit adoption.

20 Jacques Bughin and James Manyika, "How businesses are using web 2.0: A McKinsey global survey," *McKinsey Quarterly*, December 2007; Jacques Bughin and James Manyika, "Bubble or paradigm change? Assessing the global diffusion of enterprise 2.0," in Alex Koohang, Johannes Britz, and Keith Harman, eds., *Knowledge management: Research and applications,* Informing Science, 2008.

21 *Digital America: A tale of the haves and have-mores*, McKinsey Global Institute, December 2015.

Exhibit 3

AI adoption is occurring faster in more digitized sectors and across the value chain

AI index

1 The MGI Digitization Index is GDP weighted average of Europe and United States.

Source: McKinsey Global Institute AI adoption and use survey; *Digital Europe: Pushing the frontier, capturing the benefits*, McKinsey Global Institute, June 2016; *Digital America: A tale of the haves and have-mores*, McKinsey Global Institute, December 2015; McKinsey Global Institute analysis

When it comes to adopting AI, the bigger, the bolder

A stylized fact in IT literature is that large firms usually are early adopters of innovative technology, while smaller firms are more reluctant to be first movers.[22] We find the same digital divide when we look at AI: large firms have much higher rates of adoption and awareness. Across all sectors, larger firms—which we define as those with more than 500 employees—are at least 10 percent more likely than smaller firms to have adopted at least one AI technology at scale or in a core part of their business. In sectors with lower rates of AI uptake, the adoption rate of bigger companies was as much as 300 percent that of smaller companies.

Other digitization indicators reflect this fact, as highlighted in MGI's digitization work. Larger firms typically have access to more and better-structured data, and are more likely to have employees with the technical skills needed to understand the business case for AI investment and to successfully engage suppliers. Bigger firms also have an advantage because the kind of fixed-cost investment required for AI tends to generate higher returns when applied to a bigger base of costs and revenue.

Nonetheless, we find success stories among some smaller firms, too. Relative to larger companies, they can benefit from fewer issues with legacy IT systems and lower levels of organizational resistance to change. Smaller firms can also benefit from AI tools provided as a service.

Early AI adopters tend to become serial adopters

We looked at how firms deploy AI across eight different application areas and five technology systems.[23] Our results suggest that early-adopting firms are looking across multiple AI tools when they begin to adopt, rather than focusing on a particular technology. This is consistent with adoption patterns in other digital technologies.[24]

The phenomenon of multitechnology application is persistent at a sector level. Industries with high rates of adopting one technology have higher rates in adopting others. High tech and telecom, for example, report the highest rates of adoption across all five technology groups, while construction is among the lowest among all five.

However, there are anomalies. Education and healthcare are notable for being slow to adopt AI technology. In frontier sectors—those with a relatively high percentage of early adopters—

22 Kevin Zhu, Kenneth L. Kraemer, and Sean Xu, "The process of innovation assimilation by firms in different countries: A technology diffusion perspective on e-business," *Management Science*, volume 52, number 10, October 2006; Chris Forman, Avi Goldfarb, and Shane Greenstein, "The geographic dispersion of commercial Internet use," in *Rethinking rights and regulations: Institutional responses to new communication technologies*, Lorrie Faith Cranor and Steven S. Wildman, eds., MIT Press, 2003.

23 The eight technologies are: natural language processing, natural language generation, speech recognition, machine learning, decision management, virtual agents, robotics process automation, and computer vision. The five technology systems are: robotics and autonomous vehicles, computer vision, language, virtual agents, and machine learning.

24 Sanjeev Dewan, Dale Ganley, and Kenneth L. Kraemer, "Complementarities in the diffusion of personal computers and the internet: Implications for the global digital divide," *Information Systems Research*, volume 21, number 5, December 2010.

two-thirds of firms that had already adopted one of the eight AI technologies had adopted at least two others as well. In healthcare, only one-third had, with language technologies the most likely to be deployed at scale or in a core part of the business.

Users are keeping artificial intelligence close to their core

Functionally, AI technologies are finding applications across the value chain, but with some parts of the value chain getting more attention than others. For example, customer service functions such as sales and marketing, as well as operations and product development, all tend to use the most commonly cited AI applications. General and financial management, by contrast, lag well behind. A similar pattern is found in big data. The literature shows that the most frequent big data applications originate in sales and marketing functions.[25]

In general, firms queried in our survey say they tend to adopt AI technologies affecting the part of their value chain closest to the core. Operations are an important area of adoption in the automotive and assembly, and consumer packaged goods sectors, as well as utilities and resources. Operations and customer service are the most important areas for financial services. This is new. Previously, new digital technology tended to remain on the margins, away from the core of the business.

However, in line with trends in technology, we also see sectors going deeper and broader as they increase their degree of AI adoption. Leading sectors are not only more extensively deploying AI in the core parts of their value chain, they are also deploying it in more parts of their value chain.

Early adopters see AI increasing revenue while companies experimenting with AI expect lower costs

As companies become more familiar with AI, their perceptions about its benefits change. The results of survey analysis show that early AI adopters are driven to employ AI technologies in order to grow revenue and market share, and the potential for cost reduction is a secondary idea. Firms that we consider more advanced AI adopters were 27 percent more likely to report using AI to grow their market than companies only experimenting with or partially adopting AI, and 52 percent more likely to report using it to increase their market share. Experimenters, by contrast, were more focused on costs. They were 23 percent more likely than advanced AI adopters to point to labor cost reductions, and 38 percent more likely to mention non-labor cost reductions.

In other words, the more companies use and become familiar with AI, the more potential for growth they see in it. Companies with less experience tend to focus more narrowly on reducing costs.

---

25 Jacques Bughin, “Ten big lessons learned from big data analytics,” *Applied Marketing Analytics*, volume 2, number 4, 2017.

AI is not only about technical adoption, it is about enterprise acceptance

To be successful, AI adoption requires buy-in by the executive suite to generate the momentum needed to overwhelm organizational inertia.

Successful AI adopters, according to our survey, have strong executive leadership support for the new technology. Representatives of firms that have successfully deployed an AI technology at scale tended to rate C-suite support nearly twice as high as those of companies that had not adopted any AI technology. They added that strong support came not only from the CEO and IT executives—that is, chief information officer, chief digital officer, and chief technology officer—but from all other C-level officers and the board of directors as well. Successful adopters also adjusted their firm-wide strategy to become proactive toward AI.

AI's next challenge: Get users to adapt and adopt

IT industry analysts concur that the market size for AI technology will experience strong growth over the next three years. Most of the firms we surveyed expected to increase spending on AI in the coming three years, a finding echoed in other recent surveys. For example, 75 percent of the 203 executives queried in an Economist Intelligence Unit survey said AI would be "actively implemented" in their firms within three years (3 percent said it had already happened).

Expectations of how large this growth will be vary widely. Our survey documented relatively modest growth projections—only one-fifth of firms expected to increase expenditure by more than 10 percent. Industry analysts' forecasts of the compound annual growth rate ranged from just under 20 percent to nearly 63 percent, including both adoption by additional companies and increased spending within companies.[26] The actual growth rate may need to be toward the upper end of that range to meet the expectations of investors piling into the industry.

Growth will hinge on the ability of sectors and firms to overcome technical, commercial, and regulatory challenges. Our survey respondents and outside forecasters expect financial services, retail, healthcare, and advanced manufacturing to be in the AI vanguard. These are the industries where technical feasibility is relatively high (reflected in the case studies on the market today) and the business case for AI is most compelling. They are also the sectors with the highest degree of digital adoption to date—a key foundation for AI (Exhibit 4).

Technical challenges are an important differentiating factor between industries. While big tech and academia are pushing advances in the performance of the underlying technology, engineering solutions need to be worked out for specific use cases, requiring both data and talent. Industries such as financial services, high tech and telecom have generated and stored large volumes of structured data, but others, including construction and travel, lag far behind.[27]

---

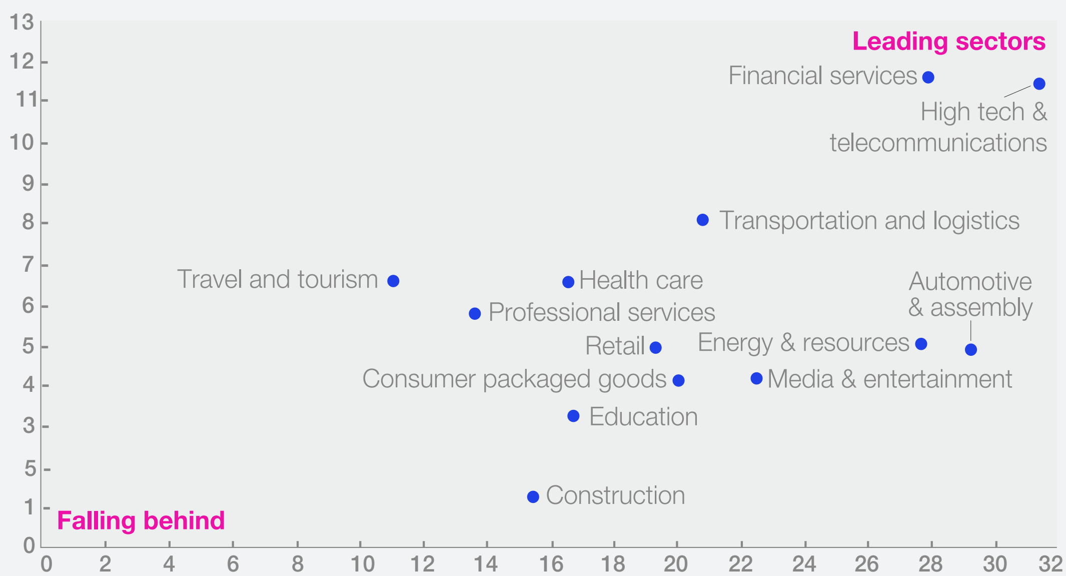
26 The full range of forecasts: BCC Research, 19.7 percent; Transparency Market Research, 36.1 percent, Tractica, 57.6 percent; IDC, 58 percent; and Markets and Markets, 62.9 percent.

27 *A future that works: Automation, employment, and productivity*, McKinsey Global Institute, January 2017.

Exhibit 4

Sectors leading in AI adoption today also intend to grow their investment the most

Future AI demand trajectory[1]
Average estimated % change in AI spending, next 3 years, weighted by firm size[2]



Current AI adoption
% of firms adopting one or more AI technology at scale or in a core part of their business, weighted by firm size2

1 Based on the midpoint of the range selected by the survey respondent.
2 Results are weighted by firm size. See Appendix B for an explanation of the weighting methodology.

Source: McKinsey Global Institute AI adoption and use survey; McKinsey Global Institute analysis

Commercial drivers also differ between sectors. Industries most likely to lead the adoption of AI technologies at scale are those with complex businesses in terms of both operations and geography, whose performance is driven by forecasting, fast and accurate decision making, or personalized customer connections. In financial services, there are clear benefits from improved accuracy and speed in AI-optimized fraud-detection systems, forecast to be a $3 billion market in 2020. In retail, there are compelling benefits from improved inventory forecasts, automated customer operations, and highly personalized marketing campaigns. Similarly, in healthcare, AI-powered diagnosis and treatment systems can both save costs and deliver better outcomes for patients.

Even where compelling commercial use cases have been engineered and are demanded by firms, regulatory and social barriers can raise the cost and slow the rate of adoption. Product liability is one such concern; it is especially troublesome for automakers and other manufacturers. Privacy considerations restrict access to data and often require it to be anonymized before it can be used in research. Ethical issues such as trained biases and algorithmic transparency remain unresolved. Preferences for a human relationship in settings such as healthcare and education will need to be navigated. Job security concerns could also limit market growth—there are already serious calls for taxes on robots.

These forces will help determine the industries that AI is likely to transform the most. However, if current trends hold, variation of adoption within industries will be even larger than between industries. We expect that large companies with the most digital experience will be the first movers because they can leverage their technical skills, digital expertise, and data resources to develop and smoothly integrate the most appropriate AI solutions.

▼▼▼

After decades of false starts, artificial intelligence is on the verge of a breakthrough, with the latest progress propelled by machine learning. Tech giants and digital natives are investing in and deploying the technology at scale, but widespread adoption among less digitally mature sectors and companies is lagging. However, the current mismatch between AI investment and adoption has not stopped people from imagining a future where AI transforms businesses and entire industries. In the next chapter, we explore the four major ways in which AI can create value across the value chain in different sectors.

Jacques Bughin is a director of the McKinsey Global Institute, **Michael Chui** and **Sree Ramaswamy** are MGI partners, and **Tera Allas** is an MGI visiting fellow; **Eric Hazan** is a senior partner in the Paris office; **Peter Dahlström** and **Nicolaus Henke** are senior partners in the London office, where **Monica Trench** is a consultant.

This article was excerpted from the McKinsey Global Institute discussion paper, Artificial Intelligence: The Next Digital Frontier? *(2017)*

Unlocking the value of analytics

Advanced analytics: Nine insights from the C-suite

Jit Kee Chin, Mikael Hagstroem, Ari Libarikian, and Khaled Rifai

Conversations with hundreds of business leaders reveal nine ways that they are—and are not—adapting to the analytics revolution.

Data and advanced analytics have arrived. The volume of available data is growing exponentially, with more added every day from billions of phones, sensors, payment systems, and cameras. Machine learning is becoming ubiquitous, but organizations are struggling to turn data into value.

The stakes are high. Those who advance furthest, fastest will have a significant competitive advantage; those who fall behind risk becoming irrelevant. Analytics cannot be the sole province of the chief information officer (CIO), as is sometimes the case. The CIO may not understand the business as a whole well enough to spot opportunities and threats, or be influential enough to ensure that the company addresses them appropriately. While the expertise the CIO brings is of course essential, business-unit leaders and CEOs must be in charge of analytics to accelerate the pace of change and to ensure intelligent investment. This is beginning to happen: McKinsey has found that more than 50 percent of CEOs consider themselves the primary leader of the analytics agenda, and that figure has been growing steadily.

With this in mind, we spoke to more than 300 top executives of major companies. Here we offer nine insights based on these conversations, and suggest actions for business leaders to take.

Analytics can create new opportunities and disrupt entire industries. But few leaders can say how

"Where do we want to be in five years as a result of advanced analytics? What are the implications to our business model, culture, portfolio mix, and value proposition?" CEOs all over the world are asking these questions—for good reason. Analytics has the potential to upend the prevailing business models in many industries, and CEOs are struggling to understand how. The need is urgent.

Beyond reorienting the existing business models, analytics leaders are also learning how to create and capitalize on new opportunities. Organizations are moving from hoarding data to sharing it. Some are pooling data as part of industry consortia, increasing their comprehensiveness and therefore their value. Product-based organizations are adding data and analytics to their offerings as value-added services. Some have gone further, charging for the analytics-enabled service rather than directly selling the product. For example, some jet-engine manufacturers now sell flight hours instead of the engines; this is only possible because sensors provide the data that help them understand usage and required maintenance.

Recommendations

There are two areas to explore. First, to understand how analytics can disrupt existing business models, set aside the time to focus on the long term. What can be learned from other industries that are farther along? What customer needs can be better met through new business models?

Second, to capture new opportunities, start with the data, analyzing what they are worth, how distinct they are, who would find them valuable, and how they can be combined with other sources to increase their value. Then, think through the business model. A simple way to get started is to conduct a market scan of the data and analytics players, as well as a competitor scan to understand what others may be doing. Identify where and how to play within this ecosystem.

Surprisingly few companies know where and how analytics can create value

Analytics create value when big data and advanced algorithms are applied to business problems to yield a solution that is measurably better than before. By identifying, sizing, prioritizing, and phasing all applicable use cases, businesses can create an analytics strategy that generates value. For example, a CEO of a global consumer-packaged-goods company told us that the application of advanced analytics and machine learning to business functions such as revenue-growth management and supply-chain optimization uncovered as much as $4 billion in benefits.

Few executives, however, have such a detailed view of value across their business units and functions. More typical is this kind of comment: "Sometimes I feel we are doing analytics for the sake of doing analytics. We need to have more clarity on what business value we are trying to create," one senior executive said. Most have experimented with a handful of use cases, but lack a comprehensive view. Even fewer have considered how analytics can create new sources of revenue. Lacking an enterprise-wide view of opportunity, business leaders struggle to make a considered business case for analytics. They may also struggle to communicate why analytics matter—and that is essential to get the organization committed to change.

Recommendations

Start a rigorous process with the executive team to decide where the most promising sources of value exist. To start, identify which functions or parts of the value chain have the most potential. For consumer-goods companies, for example, it could be product development or inventory optimization; for insurance companies, it may be risk models. Then come up with possible use cases—as many as 100 for a large company—and how new data and techniques could be applied to them. Using outside benchmarks can be useful to get a sense of how valuable a given use case might be. Finally, decide the order of priority, considering economic impact, fit with the business, feasibility, and speed.

Data science is the easy part. Getting the right data, and getting the data ready for analysis, is much more difficult

As data science enters the mainstream, commercial analytics platforms and code-sharing platforms are providing algorithm libraries and analytics tools. For most organizations, this simplifies the practical application of data science. But that still leaves the matter of what to do with it. In our conversations, we heard a familiar refrain. “The majority of our time is spent getting the data,” said a senior executive at an advanced-industries company. “Once we have that in a good place, the modeling is quick.”

Each data set is unique, and it takes time to prepare it for analyses. One major issue is that it can be difficult to agree on a “single source of truth,” because different departments often use different ways to measure the same metric. For example, the sales function may measure the volume of goods sold by transaction, while operations may measure by inventory movement. Most companies have not yet incorporated real-time data into day-to-day business processes. Many also struggle to identify what data are needed to improve competitive advantage, and therefore what they need to create. Other common challenges are implementing a unique identifier to link different data sets (such as transaction data and customer profiles) and filling in gaps to increase quality and usability.

Recommendations

The sea of data is vast and growing exponentially. To avoid drowning, executives must connect the data strategy to the analytics strategy. When exploring new data sources, it helps to have specific use cases in mind and to reflect on how data are acquired—whether through commercial vendors or via open sources. Know what data the business owns; this can become an asset to monetize. To continuously improve data quality, put in place governance and processes, and ensure that the rightful owners have direct access. Mandate good data and metadata practices and build automatic data-reconciliation processes that constantly verify that new data meet quality standards. To drive new insight, interconnect different data sets, potentially in a centralized repository (or “data lake”). Resist the temptation of complexity. Rather than building a data lake for all legacy data—a project that can take years—fill the lake gradually. Start with data required for priority use cases, and gradually add to it. Get started with what you have, and don't let perfection be the enemy of the good.

Data ownership and access need to be democratized

The most common excuse that businesses roll out for refusing to adopt counterintuitive analytics insights is that the underlying data are not valid. This claim is much more difficult to make if accountability for data quality rests with the business, and if business leaders have ready access. Successful analytics organizations give as many people as possible access to the data, while making sure there is a single source of truth, so that employees can play with them and come up with new ideas, or discard old ones that are past their prime. "The way we are thinking about eliminating the finger-pointing between business and IT on data," said the CIO of a large pharmaceutical company, "is by making data available to everyone." By doing so, a data-driven decision-making mind-set gets infused throughout the organization.

Recommendations

Design effective data governance, specifying who is responsible for data definition, creation, verification, curation, and validation—the business, IT, or the analytics center. Embrace the dual principles of business ownership and broad access. Hold the business accountable for data, even if the IT department houses and supports them. Create data-discovery platforms, such as web-based self-serve portals that allow front-line staff to easily extract data. Host data-discovery sessions to build data literacy.

Embedding analytics is as much about change management as it is about data science

Old ways of working are deeply ingrained, especially if there is an underlying distrust of analytics. Another question, then, that executives are asking is how to influence frontline staff to use the insights delivered by analytics tools to change how to make decisions. The CEO of GE, Jeff Immelt, told McKinsey: "I thought if we hired a couple thousand technology people, if we upgraded our software, things like that, that was it. I was wrong. Product managers have to be different; salespeople have to be different; on-site support has to be different."

There are some success stories. One common and essential factor is that leadership has to commit to analytics, visibly. One executive told us how the head of a business unit used analytical tools to crunch the numbers regarding stock levels. He then presented the results to the weekly leadership meeting and required each channel manager to take action.

It is also essential to integrate insights into the daily work flow. Another executive spoke about how the sales staff resisted using leads generated by the analytics model, preferring to rely on their instincts. His team was able to engineer the work flow so that the recommendation engine was "invisible": the sales team was simply presented with leads and then acted on them—successfully.

Recommendations

People buy into change when they understand it and feel they are part of it. The design of analytics solutions therefore needs to be user led and have business-process participation from the start. Have a "translator"—someone who not only understands the data science but also how it can be applied to the business—lead use-case development from start to finish. Match the talent to the task. The business identifies the opportunity, the data scientists develop the algorithm, the user-experience designers shape the user interface, the software developers run production, the process engineers reengineer work flows, and the change agents do the implementation. Develop a playbook for each use case, making sure critical adoption elements such as training and communication are not neglected. Beyond individual use cases, design a broader change program that builds analytics literacy and shifts the organization toward a data-driven culture. Organizational change management is generally well understood; it is a matter of applying these principles to analytics.

Learn to love metrics, and measure, measure, measure

"How do I know that the investment I'm making in analytics is worth it? What are the metrics? How do I attribute value to analytics versus all the other things my teams are doing?" These questions, from a senior executive at a large insurer, are typical. What's also typical is that few of the executives with whom we spoke can answer them.

If the value of analytics is not explicitly measured and then communicated, it will be difficult to build support and thus justify investment. This is not always easy, because analytics is often used to support decisions, and therefore the value cannot always be isolated from other initiatives.

In a successful measurement strategy, the metrics are detailed and logically connected to business outcomes. For each analytics use case in production, review the associated outcome metrics, and ask how they contribute to business outcomes. If the use of analytics decreases customer churn by 2 percent, how much savings does that translate into?

Recommendations

Create a dashboard that incorporates all performance indicators of interest and features automated data feeds, so that it is easy to stay on top of what is going on. Then, trust the message that the data tells. "By relying on the statistical information rather than a gut feeling," said a CEO of an investment bank, "you allow the data to lead you to be in the right place at the right time. To remain as emotionally free from the hurly-burly of the here and now is one of the only ways to succeed."

With automation and digitization, it is possible to see changes in real time, rather than waiting for the end of the month, quarter, or year. And because it is possible to measure more often, there is no excuse not to do so. Numbers only have value when they are put to work. Businesses should decide what the best cadence is, and do it.

There is no universal way to organize an analytics operating model

There are, however, two general truths. First, there should be a central function to maintain best practices and capitalize on economies of scale for hard assets. Second, accountability for value capture rests with whomever owns the bottom line. Once solutions are developed with business input, business leaders need to be held accountable for capturing the value.

What is the best operating model for analytics? The tension is between what the center of excellence (COE), a central function for data science, should be responsible for, versus where the business units are. Each model can work, if used correctly. Recent McKinsey research has found little correlation between how analytics is organized and how successful it is. What matters is that the operating model should be consistent with the business model, so that it can take advantage of the successful elements of the existing culture and practices while still promoting the cross-functional practices that any analytics effort needs to succeed.

Recommendations

Leaders should assess where the decision-making power sits in their organization—in the center or in the business units—and then design an analytics organization model that leverages the strengths of existing structures. If there is already an analytics COE, it is important to assess its effectiveness. Among the questions to consider: How fast can decisions be made? Is there sufficient business input into analytics solutions? Am I capturing the expected value from these solutions?

The talent challenge is not only to find data scientists but also to groom 'translators'

While the talent market is still tight, most CEOs we spoke to said their companies already employ data scientists. What they need is more business experts who are also proficient in analytics—translators who can spot opportunities, frame a problem, shape a solution, and champion change. "I have lots of people who speak the language of business, and I have no problem finding software engineers who speak the language of technology," one CEO told us. "But I can't find translators who speak both languages." The key is to find people who can take the numbers, and then work them for the benefit of the business.

Recommendations

Identify high performers with a quantitative background, such as statisticians and econometricians, then design a capability-building program to extend their analytics skills. The curriculum should include not only data science but also the leadership skills required to lead the identification and implementation of a use case end-to-end, and the change-management skills required to spur culture change. Make use of adult-learning principles when designing these programs, combining methods like on-the-job training, in-person learning, and online refresher courses. Consider designing formal certifications to those who successfully complete these courses. This provides recognition and creates a common language and set of standards.

The fastest way to a big idea is to cultivate a data-driven, test-and-learn culture

Every company is happy to celebrate success, which is fun and easy; but many are not so keen to communicate bad news. Many companies also have a hypothesis bias, shaping data to an existing agenda.

In many start-ups and other agile businesses, on the other hand, there is a data-driven, test-and-learn culture. Once the high-level vision is set, employees are encouraged to identify where the opportunities are, quickly develop proofs of concept, and then let the data speak to the situation. The emphasis is on generating counterintuitive insights and new ideas swiftly, testing them, and then either going ahead or tossing them out. Bad news is communicated early and without shame because mistakes are seen as sources of improvement for the next iteration. While not all parts of the organization may need to fully adopt this culture, analytics centers of excellence, as well as business units and functions that need to stay on the cutting edge, do.

Recommendations

The sandbox is a place of playful creativity in which what is built can also be quickly torn down. That is the atmosphere to aim for: provide the right tools, technology, and computer power needed to discover new features, run correlations, and perform analyses. Then, make it possible to tear it down as new information and needs supersede the old, without having to go through a lot of data security, compliance, and cleanup.

This is all part of building a culture in which data, not guesses, are brought to bear on problems, and where people are comfortable with constant change. Delivering, and hearing bad news has to be seen as part of business as usual. Set clear stage gates for investment, even while accepting that most efforts will fail, and then increase investment size as milestones are achieved. Emphasize the need for speed. "We fail more often than we succeed in analytics," noted the leader of a business unit at a consumer-goods company. "But we are trying to move more quickly in learning from failures and moving to the next iteration."

Many sectors are not getting the most out of data and analytics. Doing better requires bringing a sense of urgency to the challenge, and then a willingness to do things differently. The executives we spoke with, on the whole, understand this.

Completing a full transformation means aligning the business around a common strategic aspiration, establishing the fundamentals, and generating momentum. This typically takes two to three years. Organizations therefore have only a narrow window in which to work. Otherwise, they will fall behind—and may never catch up. As one CEO mused, "It's no longer the big fish eating the small, but the fast ones eating the slow."

Jit Kee Chin is a McKinsey alumna. **Mikael Hagstroem** is a partner in McKinsey's Charlotte office, and **Ari Libarikian** is a senior partner in the New York office, where **Khaled Rifai** is a partner.

Unlocking the value of analytics

Fueling growth through data monetization

Josh Gottlieb and Khaled Rifai

A new survey finds that many companies are launching data-focused businesses. But few have achieved significant financial impact, which requires the right combination of strategy, culture, and organization.

Results from the newest McKinsey Global Survey on data and analytics indicate that an increasing share of companies is using data and analytics to generate growth.[1] Data monetization, as a means of such growth, is still in its early days—though the results suggest that the fastest-growing companies (our high performers) are already ahead of their peers. Respondents at these companies say they are thinking more critically than others about monetizing their data, as well as using data in a greater number of ways to create value for customers and the business.[2] They are adding new services to existing offerings, developing new business models, and even directly selling data-based products or utilities.

1 The online survey was in the field from March 14–24, 2017, and garnered responses from 530 C-level executives and senior managers representing the full range of regions, industries, and company sizes. To adjust for differences in response rates, the data are weighted by the contribution of each respondent's nation to global GDP.

2 The high-performing companies are those in which respondents report annual rates of growth in organic revenue and in earnings before interest and taxes (EBIT) of 10 percent or more in the past three years.

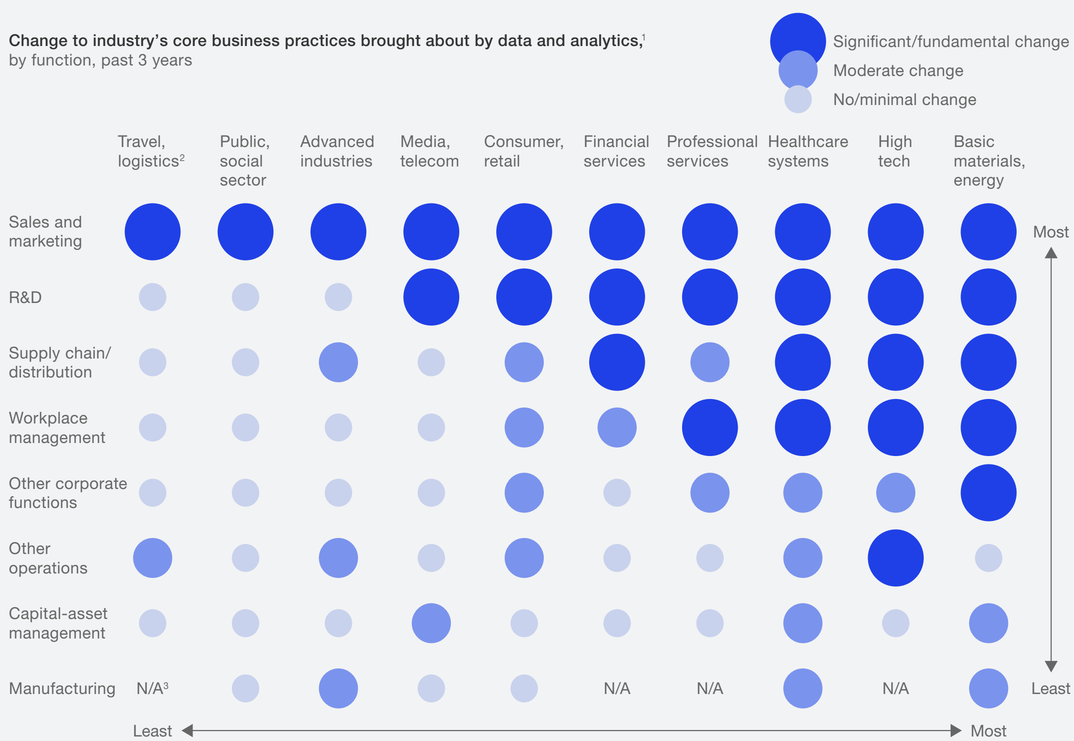
Moreover, responses from the organizations that are seeing the most impact from their data-and-analytics programs offer lessons to others striving to make the most of their data. Those companies have, according to respondents, established a strong foundation for analytics in a few ways: clear data-and-analytics strategies, better organizational design and talent-management practices, and a greater emphasis on turning new data-related insights into action.

Data and analytics are changing the way business is done

Overall, respondents say that the use of data and analytics has brought important changes to their compa-nies' core business functions. For example, nearly half of all respondents say data and analytics have significantly or fundamentally changed business practices in their sales and marketing functions, and more than one-third say the same about R&D. Across industries, respondents in high tech and in basic materials and energy report the greatest number of functions being transformed by analytics (Exhibit 1).

Exhibit 1

Data and analytics affect business practices most in the sales and marketing function and the energy and high-tech industries.

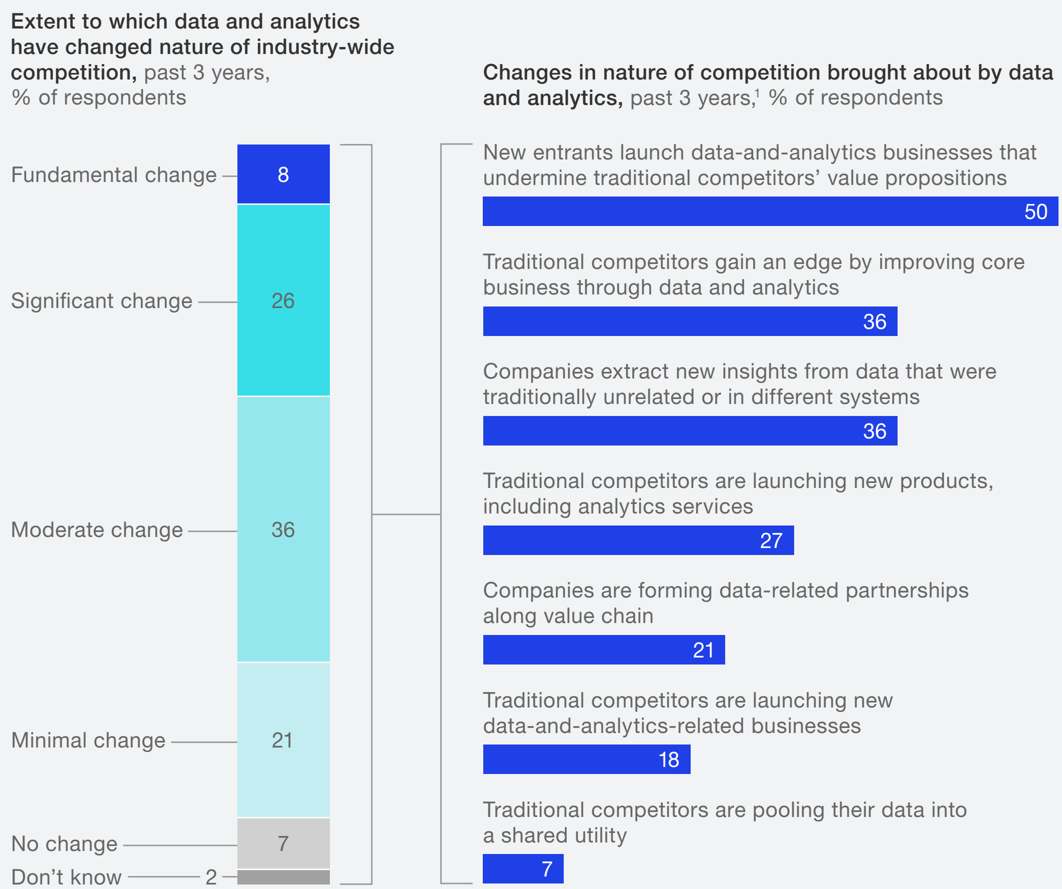


[1] Responses shown here represent the greatest degree of change (i.e., to business processes in a particular function) that at least 30% of respondents in each sector reported.
[2] In travel, transportation, and logistics, n = 36; in public and social sectors, n = 39; in advanced industries, n = 30; in media and telecom, n = 33; in consumer and retail, n = 41; in financial services, n = 85; in professional services, n = 91; in healthcare systems, n = 35; in high tech, n = 65; and in basic materials and energy, n = 48.
[3] A plurality of respondents answered "Don't know."

Data and analytics are also changing the nature of industry competition. Seventy percent of all executives report that data and analytics have caused at least moderate changes in their industries' competitive landscapes in recent years (Exhibit 2). The most common change, cited by half of respondents, is entrants launching new data-focused businesses that undermine traditional business models. Across industries, respondents report the most significant changes in high tech, media and telecom, and consumer and retail.

Exhibit 2

New data-and-analytics-related businesses and the application of data insights are changing the nature of competition.



[1]Respondents who answered "Other" and "Don't know" are not shown.

Data monetization is becoming a differentiator

Across industries, most respondents agree that the primary objective of their data-and-analytics activities is to generate new revenue. We asked about data monetization as one such way to create revenue, and the results suggest that these efforts are fairly new. Of the 41 percent of respondents whose companies have begun to monetize data, a majority say they began doing so just in the past two years.

Though nascent, monetization is already more prevalent in certain industries: more than half of the respondents in basic materials and energy, financial services, and high tech say their companies have begun monetizing data. What's more, these efforts are also proving to be a source of differentiation.

Most notably, data monetization seems to correlate with industry-leading performance. Respondents at the high-performing companies in our survey are more likely than others to say they are already monetizing data and to report that they are doing so in more ways, including adding new services to existing offerings, developing entirely new business models, and partnering with other companies in related industries to create pools of shared data (Exhibit 3). Perhaps unsurprisingly, respondents at high performers also see a top-line benefit: they are three times more likely than others to say their monetization efforts contribute more than 20 percent to company revenues.

The high performers' focus on data monetization may stem from a better ability—and greater need—to adapt to change. Compared with their peers, high-performing respondents report that data-and-analytics activities are prompting more significant changes in their core business functions. For example, respondents at high performers are at least one-third more likely to report significant or fundamental changes to business practices in areas such as supply chain, research and development, capital-asset management, and workforce management. Additionally, they are more likely to report changes in competitive pressure, whether from new entrants launching new data-related businesses, traditional rivals gaining an edge through data and analytics, or companies forming data-related partnerships along the value chain.

Get the foundations right first

Before companies can make meaningful strides with data monetization, they must first set up the fundamental building blocks of a successful data-and-analytics program.[3] We took a close look at a group of companies in which respondents report seeing the greatest business impact from analytics. The results reveal that these "analytics leaders"[4] offer important lessons as to where and how companies can strengthen their foundations, particularly in areas beyond the technical aspects of building data-and-analytics solutions (Exhibit 4).

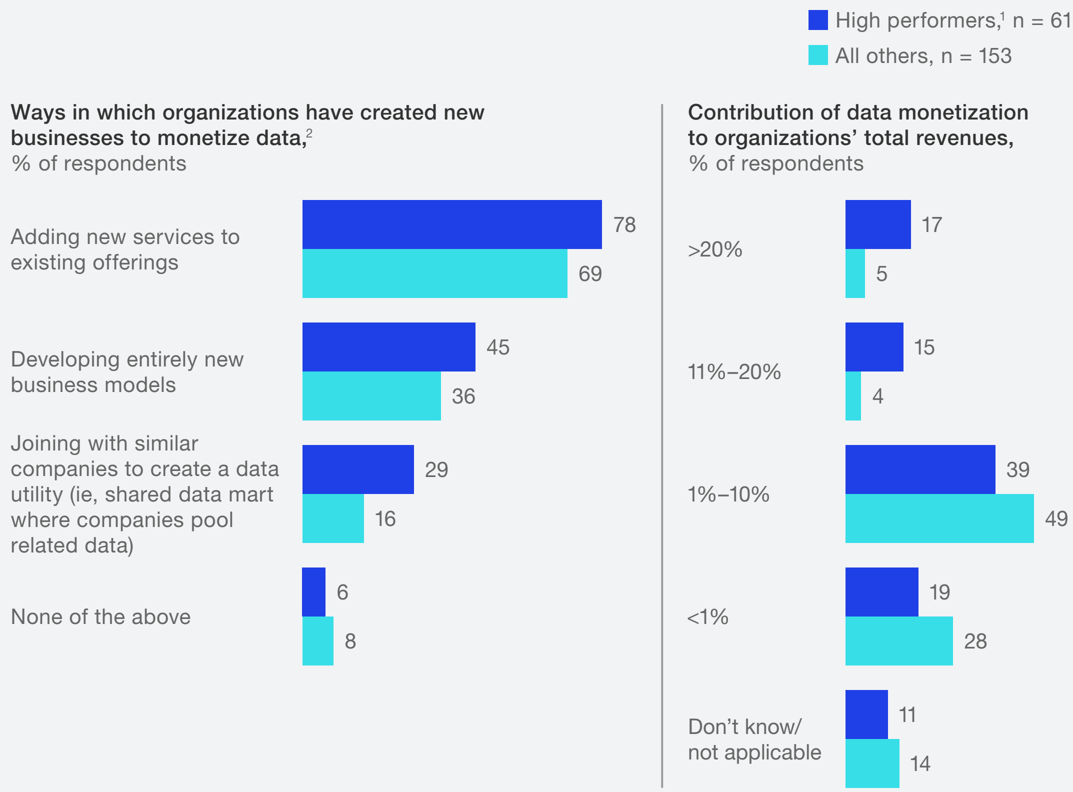
---

3 "How companies are using big data and analytics," April 2016, McKinsey.com.

4 The analytics leaders are companies that, according to respondents, have seen at least a 6 percent impact on their revenues and costs from their data-and-analytics activities in the past three years.

Exhibit 3

Compared with their peers, high performers report a greater variety of actions to monetize data—with greater revenue impact.



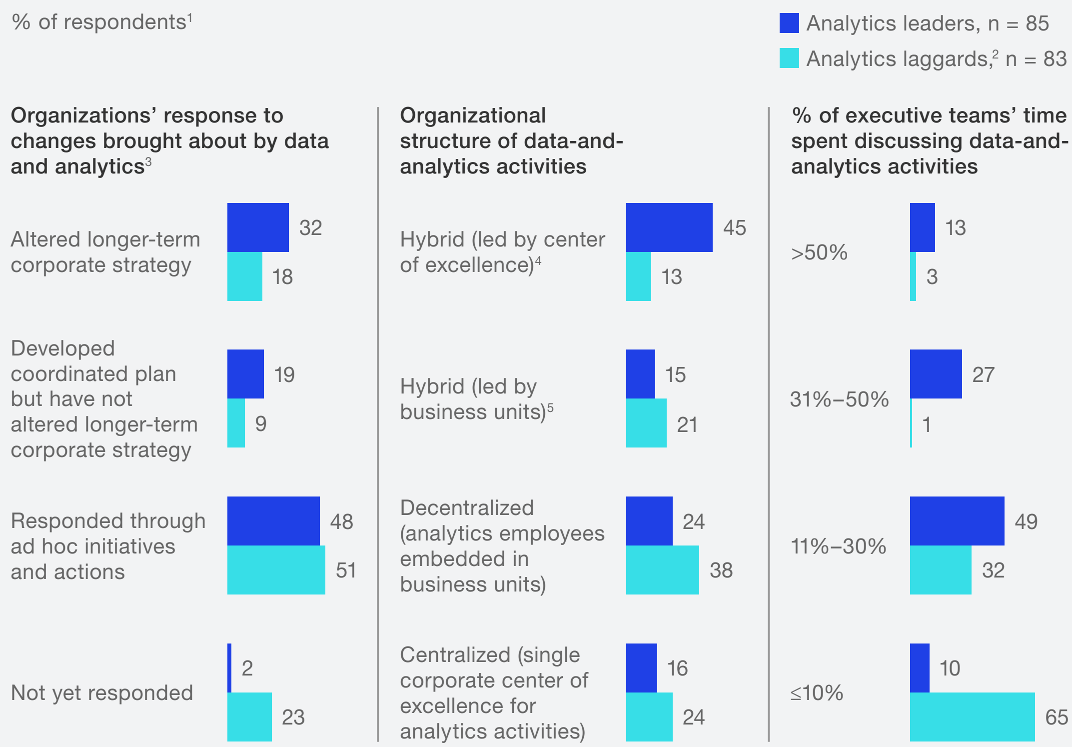
[1]High performers are organizations that, according to respondents, had annual growth rates of 10% or more for both organic revenue and earnings before interest and taxes (EBIT) over the past 3 years.
[2]Respondents who answered "Other" or "Don't know" are not shown. Question was asked only of respondents who said their organizations have already begun to monetize data.

Strategy. Many respondents report a lack of a data-and-analytics strategy at their companies, even when the need for one becomes compelling. For example, 61 percent of respondents who recognize that data and analytics have affected their core business practices say their companies either have not responded to these changes or have taken only ad hoc actions rather than develop a comprehensive, long-term strategy for analytics. In contrast, analytics leaders are nearly twice as likely as others to report enacting a long-term strategy to respond to changes in core business practices.

Organization and talent. While either a decentralized or centralized organizational model for data-and-analytics activities can work, the results suggest that a hybrid model incorporating

Exhibit 4

Analytics leaders differ from other companies in their data-and-analytics strategy, structure, and executive attention.



[1]Respondents who answered "Don't know" are not shown. Figures may not sum to 100%, because of rounding.
[2]We define a laggard as a company for which respondents say data-and-analytics activities have had less than 1% impact on (a) total revenues and (b) total costs.
[3]Question was asked only of respondents who said data and analytics brought at least minimal change to business practices in 1 or more functions; for laggards, n = 72.
[4]That is, central analytics organization sets strategy and creates tools for analytics employees in business units.
[5]That is, business units set strategy, and central analytics organization creates tools and coordinates efforts.

elements of both is much more common among the analytics leaders.[5] At the leader companies, respondents are more than three times as likely as those whose companies are struggling to see an impact from data and analytics—the laggards[6]—to say they are using a hybrid model led by a center of excellence, one of two hybrid models the survey asked about.

5 The survey asked about four types of organizational structures for data-and-analytics activities: decentralized, centralized, and hybrid models that are led by either business units or centers of excellence. Among the analytics leaders, a hybrid center-of-excellence-led model is most common (cited by 45 percent, compared with 13 percent of respondents at the analytics laggards).

6 The analytics laggards are companies that, according to respondents, have seen an impact of less than 1 percent on their revenues and costs from data and analytics in the past three years.

For all respondents—and regardless of the organizational model their companies use—attracting and retaining talent appears to be even more difficult than it was in our previous survey on the subject.[7] Nearly 60 percent of respondents now say it is harder to source talent for data-and-analytics roles than for other positions, compared with 48 percent in our previous survey. This challenge is acute even for the analytics leaders, which have a harder time than others do in finding people with both technical and domain expertise—sometimes called translators. At leading companies, 24 percent of respondents identify the translator role as their organizations' most pressing need for talent.

Leadership and culture. Successful data-and-analytics programs also require real commitment from business leaders, along with a consistent message from senior leaders on the importance and priority of these efforts. Overall, respondents report that senior-management involvement in data-and-analytics activities is the number-one contributor to reaching their objectives.[8] At the analytics leaders, senior-management practices prove the point further. Respondents at these organizations are five times more likely than those at analytics laggards to say their executive teams spend more than 20 percent of their time at high-level meetings discussing their data-and-analytics activities.

Overall, though, the survey indicates that senior-leader alignment on data-and-analytics initiatives is still not optimal at many companies. At some firms, CEOs differ from other senior leaders in their perceptions of analytics program management, organizational structure, and keys to success—a situation that creates the potential for mixed messages. For example, CEOs are much likelier than other senior executives (53 percent, compared with just 10 percent of others) to identify themselves as the leaders of their organizations' data-and-analytics agenda (Exhibit 5). CEO respondents are also more likely than others to report effectiveness at reaching data-and-analytics objectives and are less likely to view data scientists and engineers as a pressing talent need. Finally, the CEOs differ from other executives in their reasons for why their organizations have not responded to competitive or core business changes in their industries. While the others overwhelmingly cite a lack of senior-leadership commitment, CEOs are more likely to cite a lack of financial resources and uncertainty about which actions to take.

Looking ahead

Getting data monetization right requires significant effort, but it's becoming critical for staying ahead of traditional competitors and new disruptors. Based on the survey results, here are some steps executives can take to start their data-monetization efforts on the right foot:

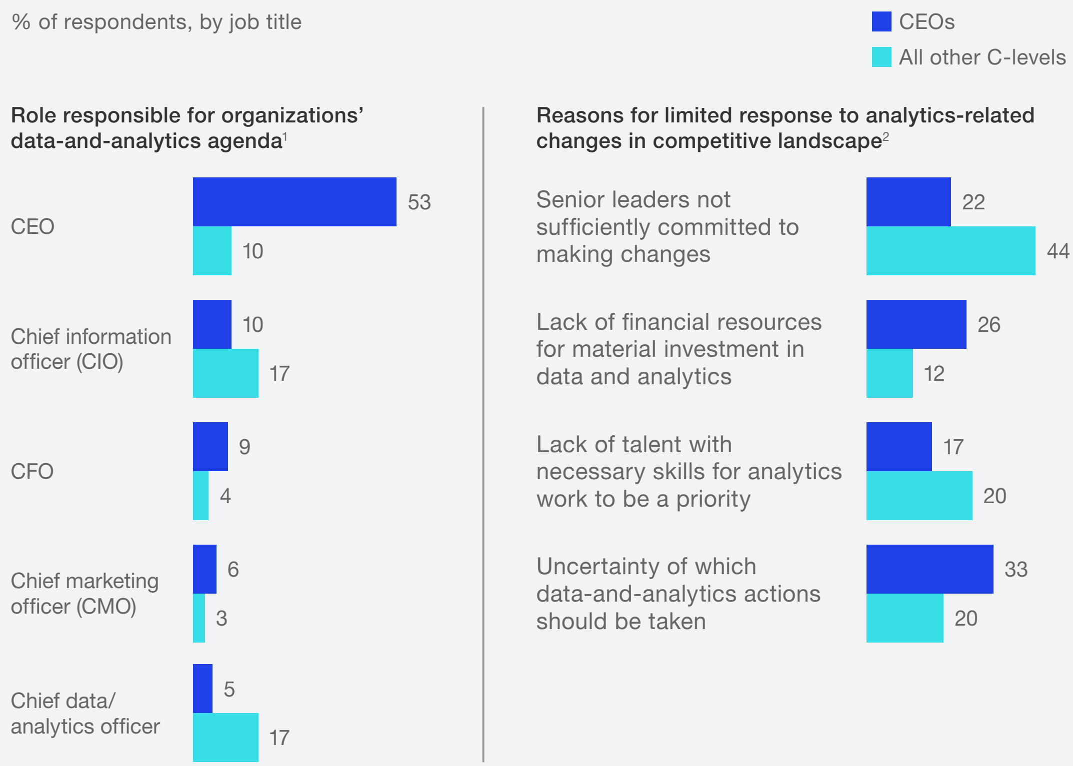
Focus on yourself first. It is nearly impossible for a company to succeed at creating externally focused data-based businesses while still struggling to get clean, consistent data that are shared internally across the organization. Before companies start down the path of monetization, they should take the time to shore up their data foundations—strategy, design, and architecture—which

7 The previous survey was in the field in September 2015.

8 Helen Mayhew, Tamim Saleh, and Simon Williams, "Making data analytics work for you—instead of the other way around," *McKinsey Quarterly*, October 2016, McKinsey.com.

Exhibit 5

CEOs often see themselves as heads of the data-and-analytics agenda and cite different reasons for not pursuing analytics activities.



[1]Out of 9 roles that were offered as answer choices; roles are arranged in descending order, based on CEO responses to the question. For CEOs, n = 269; for all other C-level respondents, n = 182.
[2]Respondents who answered "don't know" are not shown, and question was asked only of respondents who said their organizations have not responded to changes in industry-wide competition due to data and analytics, or whose organizations have responded through ad hoc initiatives. For CEOs, n = 144; for all other C-level respondents, n = 97.

will help them build the business case and technical platform they need to monetize data effectively. Putting their data to work for internal use cases, such as improving decision making or optimizing operations, can also serve as a testing ground for their data foundations as well as for the data-monetization models of new data-based businesses.

Look outside for innovation. Once companies' data-and-analytics foundations are in place, they may still find that the most innovative solutions can best be sourced externally by partnering with others in the data ecosystem.[9] Such partners include analytics companies that can supplement the organization's existing capabilities, platform providers that host tools or solutions, and data providers

9 "As sector borders dissolve, new business ecosystems emerge," *McKinsey Quarterly*, October 2017, McKinsey.com.

that can help the organization gain access to unique data sets. Companies can even work with suppliers, customers, or their industry peers to augment and enrich existing data; they can then offer those data as unique add-ons to existing products or services, or sell the data as part of an entirely new business.

Commit to an end-to-end transformation, and get the business involved. Even as data monetization gains steam, many companies are still struggling to drive major business impact. In our experience, this happens for two reasons: failure to make the wholesale changes required to enter new markets, and a lack of partnership between the business and IT. For a transformation, such changes could involve the reconfiguration of operating models and core business functions (from product development to marketing), worker-reskilling programs, and change-management programs aimed at shifting organizational culture, mind-sets, and behaviors. These sorts of substantial efforts require full commitment from the C-suite, which must communicate to senior managers—in both business units and technology centers—the priority of a given initiative or program and the need to dedicate adequate time, human capital, and financial resources to make it succeed. Many companies also struggle with data monetization—and, in particular, finding the right strategy—when they delegate all data-and-analytics efforts to IT. In reality, efforts to monetize data are more effective when they are business led and focused on the most valuable use cases.

The contributors to the development and analysis of this survey include **Josh Gottlieb**, a specialist in McKinsey's Atlanta office, and **Khaled Rifai**, a partner in the New York office.

Building foundations

A smarter way to jump into data lakes

Mikael Hagstroem, Matthias Roggendorf, Tamim Saleh, and Jason Sharma

An agile approach to data-lake development can help companies launch analytics programs quickly and establish a data-friendly culture for the long term.

Increases in computer processing power, cloud-storage capacity and usage, and network connectivity are turning the current flood of data in most companies into a tidal wave—an endless flow of detailed information about customers' personal profiles, sales data, product specifications, process steps, and so on. The data arrive in all formats and from a range of sources, including Internet-of-Things devices, social-media sites, sales systems, and internal-collaboration systems.

Despite an increase in the number of tools and technologies designed to ease the collection, storage, and assessment of critical business information, many companies are still unsure how best to handle these data. Business and IT leaders have told us they remain overwhelmed by the sheer volume and variety of data at their disposal, the speed at which information is traversing internal and external networks, and the cost of managing this wealth of business intelligence. Increasingly, they are also being charged with an even more complicated task: harnessing meaningful insights from this wealth of business information.

These executives must expand their data-management infrastructures massively and quickly. An emerging class of data-management technologies holds significant promise in this regard: data lakes. These storage platforms are designed to hold, process, and analyze structured and unstructured data.[1] They are typically used in conjunction with traditional enterprise data warehouses (EDWs), but in general, they cost less to operate than EDWs. Cost savings result because companies can use affordable, easy-to-obtain hardware and because data sets do not need to be indexed and prepped for storage at the time of induction. Data are held in their native formats and reconfigured only when needed. Relational databases may also need to be managed as part of the data-lake platform, but only to ease end users' ability to access some data sources.

There is a lot for companies to like about data lakes. Because data are loaded in "raw" formats rather than preconfigured as they enter company systems, they can be used in ways that go beyond just basic capture. For instance, data scientists who may not know exactly what they are looking for can find and access data quickly, regardless of format. Indeed, a well-maintained and governed "raw data zone" can be a gold mine for data scientists seeking to establish a robust advanced-analytics program. And as companies extend their use of data lakes beyond just small pilot projects, they may be able to establish "self-service" options for business users in which they could generate their own data analyses and reports.

However, it can be time consuming and complicated to integrate data lakes with other elements of the technology architecture, establish appropriate rules for company-wide use of data lakes, and identify the supporting products, talent, and capabilities needed to deploy data lakes and realize significant business benefits from them. For instance, companies typically lack expertise in certain data-management approaches and need to find staffers who are fluent in emerging data-flow technologies such as Flume and Spark.

In many cases, companies are slowing themselves down. They are falling back on tried-and-true methods for updating technology architectures—for instance, engaging in long, drawn-out internal discussions about optimal designs, products, and vendors and holding off on building a data-lake solution until they have one that is just right. In the meantime, opportunities to deploy advanced-analytics programs that will support digital sales and marketing and new-product development simply pass them by.

Companies should instead apply an agile approach to their design and rollout of data lakes—piloting a range of technologies and management approaches and testing and refining them before getting to optimal processes for data storage and access. The companies that do can keep up with rapidly changing regulatory and compliance standards for data—for instance, the European Union's General Data Protection Regulation, which is slated to take effect in May 2018. Perhaps more important, they can bring analytics-driven insights to market much faster than their

1 "Structured" data (such as an Excel spreadsheet) are well organized and therefore easily identified by search algorithms; "unstructured" data (such as an audio file) are less organized and therefore less likely to be responsive to search algorithms.

competitors while significantly reducing the cost and complexity of managing their data architecture.

Stages of data-lake development

Companies generally go through the following four stages of development when building and integrating data lakes within their existing technology architectures (exhibit):

- **Landing and raw-data zone.** At the first level, the data lake is built separate from core IT systems and serves as a low-cost, scalable, "pure capture" environment. The data lake

Exhibit

Companies may go through any or all of these four stages of building and integrating data lakes within technology architectures.

Stage 1: Landing zone for raw data	Stage 2: Data-science environment	Stage 3: Offload for data warehouses	Stage 4: Critical component of data operations
Data lake is a low-cost, scalable, "pure capture" environment	Data lake is actively used as a platform for experiments	Data lake is integrated with existing enterprise data warehouses (EDWs).	Data lake is a core part of the data infrastructure.
• Data lake is built separate from core IT systems. • Data are stored in raw formats. • Internal data can be easily complemented with or enriched by external sources of data.	• Data lake becomes a test-and-learn environment. • Data scientists analyze unaltered data and build prototypes for analytics programs. • IT organization deploys "just enough" data governance.	• High-intensity, mass-extraction tasks remain in EDWs … • … but large, more detailed sets of data are pushed to the data lake, in the process, easing storage and cost constraints. • Data lake can be used for "needle in a haystack" searches or other tasks that do not require traditional indexing.	• Data lake can now replace operational data stores and enable "data-as-a-service" options. • Businesses can better handle computing-intensive tasks, such as machine-learning programs. • Data-intensive applications or application programming interfaces may be built on top of the data lake. • IT organization deploys "strong" data governance.

serves as a thin data-management layer within the company's technology stack that allows raw data to be stored indefinitely before being prepared for use in computing environments. Organizations can deploy the data lake with minimal effects on the existing architecture. Strong governance, including rigorous tagging and classification of data, is required during this early phase if companies wish to avoid creating a data swamp.

- **Data-science environment.** At this next level, organizations may start to more actively use the data lake as a platform for experimentation. Data scientists have easy, rapid access to data—and can focus more on running experiments with data and analyzing data, rather than focusing solely on data collection and acquisition. In this sandbox, they can work with unaltered data to build prototypes for analytics programs. They may deploy a range of open-source and commercial tools alongside the data lake to create the required test beds.

- **Offload for data warehouses.** At the next level, data lakes are starting to be integrated with existing EDWs. Taking advantage of the low storage costs associated with a data lake, companies can house "cold" (rarely used, dormant, or inactive) data. They can use these data to generate insights without pushing or exceeding storage limitations, or without having to dramatically increase the size of traditional data warehouses. Meanwhile, companies can keep high-intensity extraction of relational data in existing EDWs, which have the power to handle them. They can migrate lower-intensity extraction and transformation tasks to the data lake—for instance, a "needle in a haystack" type of search in which data scientists need to sweep databases for queries not supported by traditional index structures.

- **Critical component of data operations.** Once companies get to this stage of rollout and development, it is very likely that much of the information that flows through the company is going through the data lake. The data lake becomes a core part of the data infrastructure, replacing existing data marts or operational data stores and enabling the provision of data as a service. Businesses can take full advantage of the distributed nature of data-lake technology as well as its ability to handle computing-intensive tasks, such as those required to conduct advanced analytics or to deploy machine-learning programs. Some companies may decide to build data-intensive applications on top of the data lake—for instance, a performance-management dashboard. Or they may implement application programming interfaces so they can seamlessly combine insights gained from data-lake resources with insights gained from other applications.

The time and capabilities required for companies to grow their data lakes from simple landing zones to critical components of the data infrastructure will vary depending on companies' objectives and starting points. At each stage of development, companies need to examine complicated questions relating to the size and variety of their data sets, their existing capabilities in data management, the level of big data expertise in their business units, and product knowledge in the IT organization. For instance, how sophisticated are analytics tools in the current environment? Is the company using traditional development tools and methodologies, or newer ones? How many concurrent data users does the company typically require? Are workloads managed dynamically? How quickly do end users need access to data? At various points in the data-lake development process, companies can get mired in these details and lose momentum; leaders in the IT organization or the business units inevitably fan out to tackle other "urgent" projects.

The data lake's journey from "science project" to fully integrated component of the data infrastructure can be accelerated, however, when IT and business leaders come together to answer these and other questions under an agile development model. In our experience, an agile approach can help companies realize advantages from their data lakes within months rather than years. Quick wins and evidence of near-term impact can go a long way toward keeping IT and business leaders engaged and focused on data-management issues—thereby limiting the need for future rework and endless tweaking of protocols associated with populating, managing, and accessing the data lake. An agile approach can put IT and business leaders on the same page. Such collaboration is critical not just for determining a technical path forward for the data lake but also for establishing a data-friendly work environment and seizing new business opportunities based on insights from data.

Building a data lake: An agile approach

Most organizations understand the need for agile methodologies in the context of software development. Fewer have applied agile in the context of data management. Typically, the IT organization takes the lead on vetting potential technology options and approaches to building data lakes, with little input from the business units. Under an agile approach, IT and business leaders jointly outline and address relevant technology and design questions. For instance, will the data lake be built using a turnkey solution, or will it be hosted in the cloud (using private, public, or hybrid off-site servers)? How will the data lake be populated—that is, which data sets will flow into the lake and when? Ideally, the population of the data lake should be based on the highest-priority business uses and done in waves, as opposed to a massive one-time effort to connect all relevant data streams within the data lake.

Indeed, the most successful early adopters are designing their data lakes using a "business back" approach, rather than considering technology factors first. They are identifying the scenarios in which business units could gain the most value from the data lake and then factoring those scenarios into the design (or redesign) of the storage solution and rollout decisions. Companies are then incrementally populating the data lake with data for specific groups or use cases, as needed. And rather than going all in on one designated solution, companies are piloting two or three final candidates from different providers to assess the real-world performance, ease of integration, and scalability of their offerings.

This agile approach to rollout can ensure that performance or implementation challenges will be caught early. It incorporates feedback from the business units. It also leaves room for agile development teams to tinker with processes and data-governance protocols as the data lake fills up, analytics and storage technologies change, and business requirements evolve.

As data lakes move from being pilot projects to core elements of the data architecture, business and technology leaders will need to reconsider their governance strategies. Specifically, they must learn to balance the rigidity of traditional data oversight against the need for flexibility as data are rapidly collected and used in a digital world. Under an agile approach to governance, businesses can apply sufficient oversight as new sources enter the data lake, avoiding some of the more rigid engineering practices required in traditional data warehouses and then refining rules and processes as business requirements dictate to get to an optimal solution. For instance, data scientists might be given free rein to explore data, even as business cases for certain categories of data are still being identified. Meanwhile, frontline users might face stricter controls until use cases are more firmly established.

At the very least, however, companies should designate certain individuals as owners of data sets and processes, so that responsibilities are clear and decisions about data sources and access rights can be made quickly. Because data are not being structured up front, companies will also want to capture and store metadata on all the data sources flowing into the lake (either within the lake itself or in a separate registry) and maintain a central data catalog for all stakeholders. Additionally, companies may need to reconfigure access rights as they iterate on data-management protocols—keeping in mind regulatory requirements and privacy issues related to holding personally identifiable information. Data owners must communicate these access rights to all relevant stakeholders.

Transformation at a global bank

Let's consider how a global bank applied agile principles to its development of a data lake. The bank had been struggling with several critical data challenges: low-quality business information, lack of specialists to manage different data sets arriving in different formats, aging data-warehouse technologies, and more than 1,000 data sources. The systems were kludgy. Incoming data sets had to be structured before they could be entered into four data-warehouse layers (output delivery, normal form, subject layer, and app layer) and before any usable reports could be created.

Outside of these technical challenges, business and IT leaders at the bank were not working collaboratively, which exacerbated the company's data problems. Data were being stored in isolated systems, so critical business information often remained trapped. But requests for access to certain data sets were slow to get a response because of poor coordination and communication across business units and IT operations. Data management was seen as "IT's job"; business leaders held the topic at arm's length and thus struggled to articulate their data needs.

Senior leaders at the bank were concerned about losing customers, in part due to the company's inability to manage data adroitly. They decided to experiment with data-lake technologies to try to ease the extraction, structuring, and delivery of data sets. Seeking to work as quickly as its software developers, the company used an agile development model and rolled out the data-lake project in phases.

Senior leaders convened an agile data team involving subject-matter experts from the business units and from the IT organization to consider the business impact of and use cases for improved data quality and access before determining which areas of the company would have initial access to the data lake.

The agile data team conducted in-depth interviews with business users to identify pain points and opportunities in existing data-management practices. The team's plan was to release waves of new data services and applications in four-month windows—implementing new data-management tools, developing data-delivery services with the business units, and refining processes based on customers' feedback. Within months of the initial launch of the agile data project, the bank was able to load data relevant to particular business use cases into a common environment and identify the critical data elements required for providing services to the business units.

Success in high-profile areas of the business enabled the bank to extend the usage of the data lake to other areas in subsequent months. The shift from structuring all the data up front to documenting a back-end process only for utilized data was significant. The bank was able to break down data silos; information from systems could now be found in one place, and employees were able to access multiple forms of data (demographic, geographic, social media, and so on) to gain a 360-degree view of customers. Collaboration between the business units and the IT group increased, as did employees' and customers' satisfaction scores.

▼▼▼

More and more companies are experimenting with data lakes, hoping to capture inherent advantages in information streams that are readily accessible regardless of platform and business case and that cost less to store than do data in traditional warehouses. As with any deployment of new technology, however, companies will need to reimagine systems, processes, and governance models. There will be inevitable questions about security protocols, talent pools, and the construction of enterprise architecture that ensures flexibility not just within technology stacks but also within business capabilities. Our experience suggests that an agile approach to the implementation of data lakes can help companies climb the learning curve quickly and effectively.

Mikael Hagstroem is a partner in McKinsey's Charlotte office, **Matthias Roggendorf** is a senior expert in the Berlin office, **Tamim Saleh** is a senior partner in the London office, and **Jason Sharma** is an associate partner in the Silicon Valley office.

The authors wish to thank Prasoon Sharma for his contributions to this article.

Building foundations

Why you need a digital data architecture to build a sustainable digital business

Sven Blumberg, Oliver Bossert, Hagen Grabenhorst, and Henning Soller

Companies that succeed at meeting their analytics objectives let business goals drive the technology. Here's how they structure a data architecture that works.

Data architecture has been consistently identified by CXOs as a top challenge to preparing for digitizing business. Leveraging our experience across industries, we have consistently found that the difference between companies that use data effectively and those that do not—that is, between leaders and laggards—translates to a 1 percent margin improvement for leaders. In the apparel sector, for instance, data-driven companies have doubled their EBIT margin as compared to their more traditional peers.

Using data effectively requires the right data architecture, built on a foundation of business requirements. However, most companies take a technology-first approach, building major platforms while focusing too little on killer use cases.

Many businesses, seeing digital opportunities (and digital competition) in their sectors, rush to invest without a considered, holistic data strategy. They either focus on the technologies alone or address immediate, distinct use cases without considering the mid- to long-term creation of sustainable capabilities. This goes some way toward explaining why a 2017 McKinsey Global Survey found that only half of responding executives report even moderate effectiveness at meeting their analytics objectives. The survey found the second-largest challenge companies face (after constructing a strategy to pursue data and analytics) is designing data architecture and technology infrastructure that effectively support data-and-analytics activities at scale. We found that eight out of ten companies embark on digital data enablement by making their IT departments responsible for the data transformation—with very grand implementation programs—and a small set of business use cases.

This strategy is quite different from that employed by next-generation digital leaders, who typically embark on transformation from a business perspective and implement supporting technologies as needed. Doing the technology first produces more problems than successes, including:

- **Redundant and inconsistent data storage.** Only two in ten banks we've worked with have established a common enterprise data warehouse, which is essential for creating a single source of truth for financial and customer data.

- **Overlapping functionality.** Every bank we've worked with has at least one business function supported by three different technological systems.

- **A lack of sustainability.** The solutions at which financial institutions typically arrive are often quick fixes that ignore the enterprises' larger aspirations for datafication. For example, one insurance company extracted and replicated data from its warehouse each time it was needed rather than building data architecture that would allow it to store each customer element only once, thereby reducing costs and eliminating inefficiencies.

These problems have real business consequences. Meeting leading-edge business requirements, such as real-time customer and decision support, and large-scale analytics requires the integration of traditional data warehousing with new technologies.

The two-speed data-architecture imperative

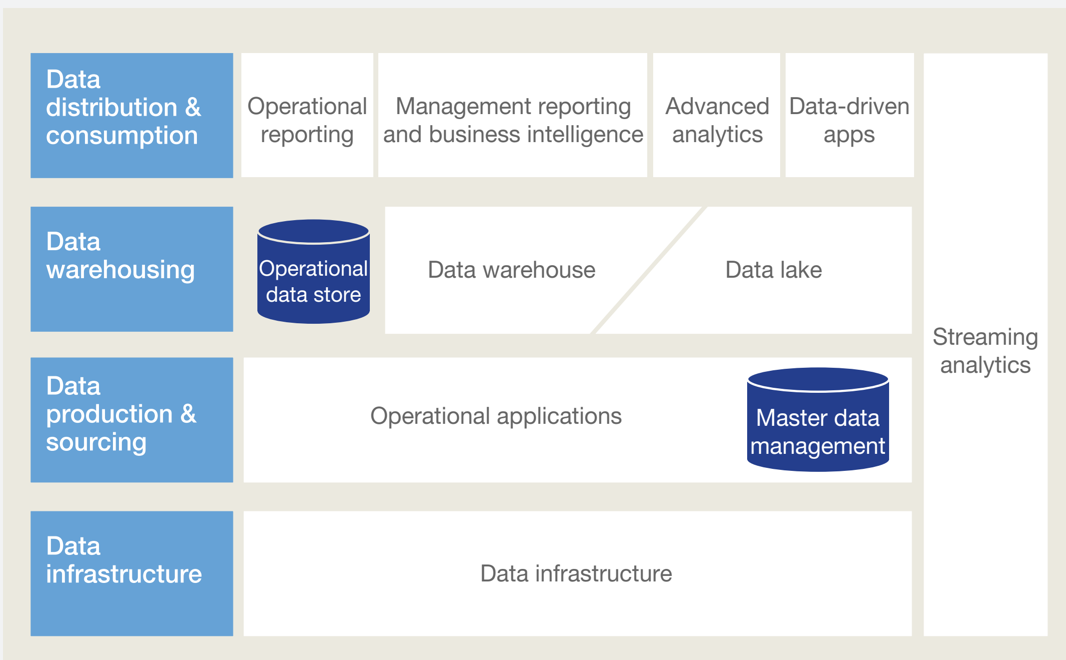
Today, enterprises must cope with increasingly large and complex data volumes (worldwide, data storage doubles every two years) coming from diverse sources in a wide variety of formats that traditional data infrastructures struggle, and most often fail, to operationalize. Developing new business capabilities—such as individual pricing for customers based on real-time profitability, as some insurance companies have done, automating credit decisions that lead to improved outcomes for banks and greater customer satisfaction, or running automated, more cost-effective strategic marketing campaigns as we've seen in the chemicals sector—demands new ways of managing data. This does not mean, however, that legacy data and IT infrastructures must be trashed, or that new capabilities need to be bolted on. It does mean that the traditional data warehouse, through which the organization gains stability and financial transparency, must be scaled down and integrated with the high-speed transactional architecture that gives the organization the capability to support new products and services (as well as real-time reporting).

This is the two-speed principle.

This new, complex technical environment requires companies to closely examine business use cases before making costly technology decisions, such as needlessly ripping out and replacing legacy architectures. Instead, it is preferable to use a capability-oriented reconceptualization of data management as an enabler of digital applications and processes (Exhibit).

Exhibit

Best-practice reference-data architecture



To implement an end-to-end digital data architecture, an enterprise needs first to develop a point of view on its current and, if possible, future business requirements, sketch its desired, flexible data-management architecture, and create a roadmap for implementation. To begin, one must identify the key business use cases.

To do this, we recommend a thorough review of best-practice use cases across industries that address common value drivers (financial transparency, customer satisfaction, rapid product development, real-time operational reporting, and so on). Then, the company should compare those use cases with its market position and strategic direction, prioritizing those that best reflect the company's situation and aspirations. Once those reference use cases are identified, the company can begin to define target data-architecture capabilities. In this process, the business leads and technology follows.

The high-level structure in the exhibit above represents a layered data architecture that has been applied successfully by many organizations across many industries, especially in finance. It extends to accommodate new digital capabilities such as collecting and analyzing unstructured data, enabling real-time data processing, and streaming analytics.

The exhibit shows a reference architecture that combines both the traditional requirements of financial transparency via a data warehouse and the capability to support advanced analytics and big data. In a phrase, it's a two-speed approach.

The two-speed architecture adheres to three core principles:

1. A limited number of components with a clear demarcation of capabilities to manage complexity while providing the required functionalities, such as advanced analytics and operational reporting

2. Layers that enable the transparent management of data flows and provide a single source of truth to protect against silos and data inconsistencies (through the data warehouse, which models, integrates, and consolidates data from various sources)

3. Integration of state-of-the-art solutions with traditional components, such as the data warehouse, to satisfy such new requirements as real-time processing, and an operational data store (ODS) based on new database technologies

We have used this model to:

- Help clients think through and evaluate their options on an architectural level before discussing concrete technical solutions.

- Map technology components against capabilities to manage and avoid redundancies while identifying gaps.

- Create plans for stepwise transformations driven by business value while limiting business disruption.

Getting physical with digital

For example, one of the largest banks in Scandinavia, understanding the business potential of advanced analytics, big data, and better data management to improve fraud detection and prevention, ATM location, and other initiatives, was eager to begin its digital data journey. It was facing intense competition and was considering making a massive, multimillion-dollar investment in its IT and data architecture.

A lot was riding on what the bank decided to invest in, where it decided to invest it, and how.

It began by identifying key use cases that reflected the organization's most compelling strategic requirements: improved fraud detection, optimized location and allocation of branches, and more granular customer segmentation.

Based on this determination, we helped the bank outline a target architecture, founded on the best-practice reference model, that would enable the capabilities the bank desired and assess available solutions. Instead of ripping out its entire IT infrastructure, the bank decided to add a single Hadoop solution that allowed for storage and distributed processing of the bank's extremely large and frequently unstructured data sets across thousands of individual machines. This was especially useful in scaling the bank's high-frequency requirements for its online fraud-detection processes.

For branch location, allocation, and optimization, a Hadoop data lake (a management platform that processes flat, nonrelational data) used the bank's geospatial and population-growth data to determine where best to locate new branches and ATM machines. To improve its customer segmentation, the bank tested a new customer algorithm on the Hadoop database before rolling it out on its legacy data warehouse. This eliminated the typically costly and time-consuming back-and-forth process of develop, pilot, assess, validate, tweak, and pilot again that characterizes traditional data developments.

In this way, the bank achieved its primary business goals. It added new, differentiating capabilities, such as real-time analytics, and created real enterprise value with a relatively small technology investment, not the massive one originally contemplated. This was achieved by deciding what to invest in, where to invest it, and how—before buying systems and software that might not have served it nearly as well. Crucially, instead of first buying the technology, the bank built an in-house analytics team, skimming off the cream of the local talent in the process.

Today, the bank is considered the leader in financial analytics in its market and sells analytics services to other financial institutions.

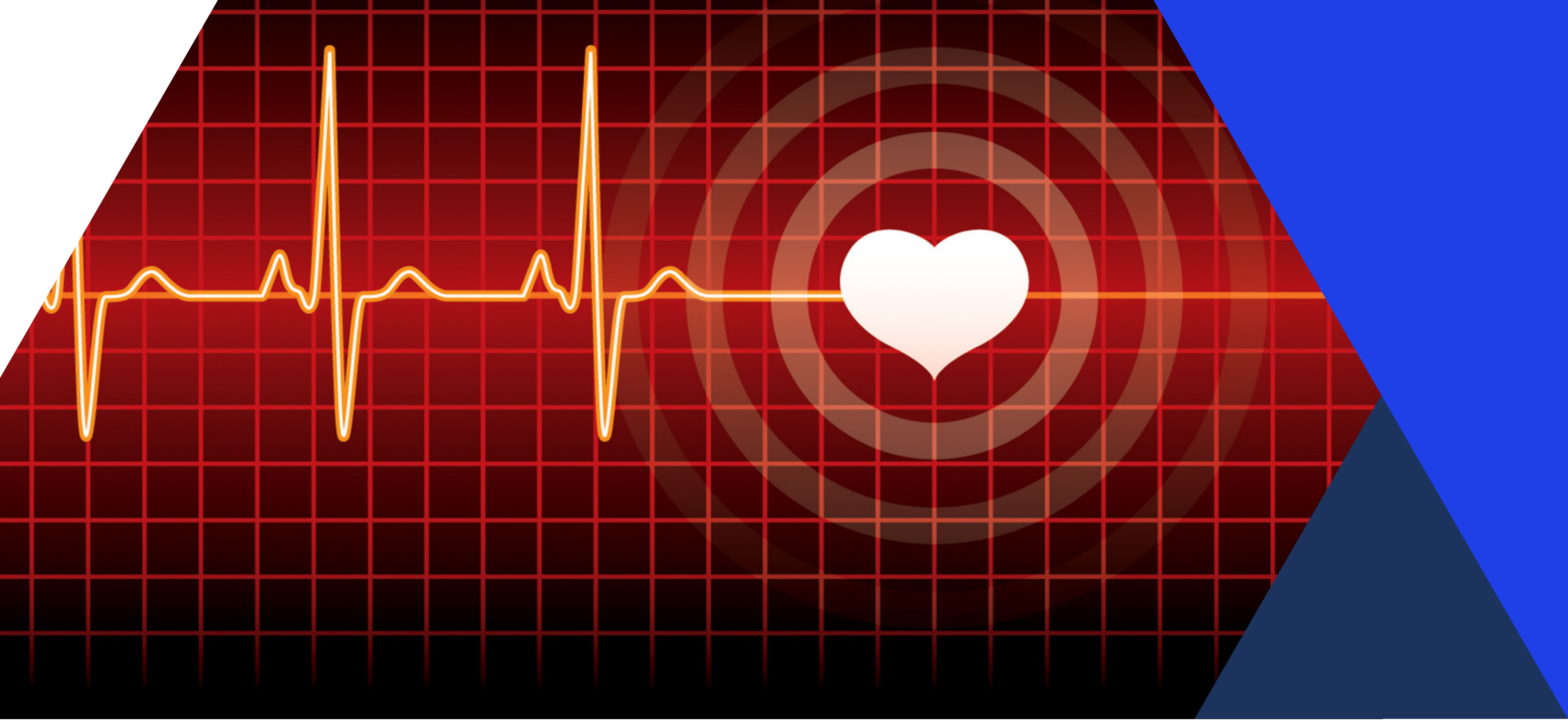
The bank knew that the time was ripe to get serious about digital transformation, made it a priority, and in doing so achieved what may well be an enduring competitive advantage, all without disrupting its business with a big-bang technological transformation. It started with a clear view of its business goals, kept them front and center, and created a two-speed data architecture that worked.

The lesson here is that for many companies, it is both doable and cost-effective to add analytics capabilities to an existing IT environment. But that requires a sound data architecture and a well-grounded approach to data management.

Sven Blumberg is a partner in McKinsey's Düsseldorf office, and **Oliver Bossert** is a senior expert in the Frankfurt office, where **Hagen Grabenhorst** is a consultant and **Henning Soller** is an associate partner.

PART 2

Analytics grows across the organization



Marketing & Sales

The heartbeat of modern marketing: Data activation & personalization

Julien Boudet, Brian Gregg, Jason Heller, and Caroline Tufft

Technology has finally advanced to the point where marketers can use real-time data in a way that is both meaningful to customers and profitable for companies.

We've come a long way from "People who bought this, also bought that."

Consider the experience of a representative customer we'll call Jane. An affluent, married mom and homeowner, Jane shops at a national clothing retailer online, in the store, and occasionally via the app. When visiting the retailer's website in search of yoga pants, she finds style choices based on previous purchases, the purchases of customers with profiles similar to hers, and the styles of yoga pants most frequently purchased on weekends. She adds one of the offered yoga pants to her shopping cart and checks out.

With the exception of a follow-up email, most interactions with the customer stop there. But here's what this example looks like when we activate Jane's data. Three days after her online purchase, the retailer sends Jane a health-themed email. Intrigued, she clicks the link and watches a video about raising healthy kids. One week later, she receives an iPhone message nudging her to use the store's mobile app to unlock a 15 percent one-day discount on workout equipment. Though she has never bought such items at this retailer, Jane takes advantage of the offer and purchases a new sports bag. What began as a simple task of buying yoga pants ended up being a much more engaged experience

Such data-activated marketing based on a person's real-time needs, interests, and behaviors represents an important part of the new horizon of growth. It can boost total sales by 15 to 20 percent, and digital sales even more while significantly improving the ROI on marketing spend across marketing channels: from websites and mobile apps to—in the not-too-distant future—VR headsets and connected cars.

Customer-data platform: Solving the ongoing challenge of true personalization

Companies regularly experiment with testing the impact of varied customer experiences, but they do it in isolation. When they do try to scale, they smack against the challenge of understanding what to prioritize. Going back to Jane, do marketers target her as a mom, a yoga enthusiast, or a homeowner? What happens when tests are running against all three segments? Is she part of a new microsegment that combines attributes and signals across all three segments?

This is a challenge that has continued to plague marketers, despite the promise of solutions such as customer-relationship management (CRM), master-data management (MDM), and marketing-resource management (MRM). These solutions can help companies consolidate and streamline data, manage segmentation, organize workflow, and improve customer relationships. But they don't take full advantage of digital signals customers provide. Instead, they rely on antiquated "list pulls," basic segmentation, and campaigns, all of which lack the automated decision making, adaptive modeling, and nimble data utilization to scale personalized interactions.

Enter the Customer Data Platform (CDP)—a data discovery and "decisioning" (i.e., automated decision making) platform. The CDP makes it possible for marketers to scale data-driven customer interactions in real time. And while CDP hasn't really broken into the Gartner Magic Quadrant or Forrester Wave, it is gradually becoming an industry-standard concept, with a small but growing cadre of third-party platforms emerging that will soon shape the category.

Four steps to effectively activate your data

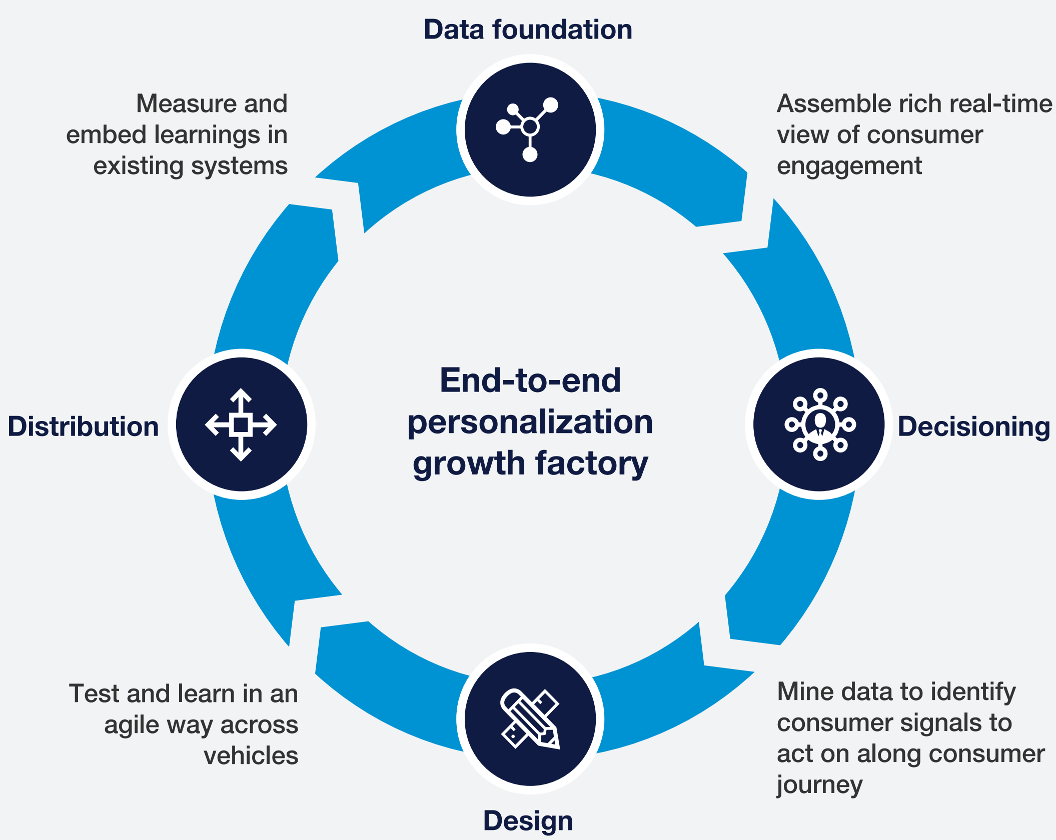
Incorporating a CDP into your organization—whether piggybacking on an existing master data-management or customer-relationship-management system or starting from scratch—requires mastery of four areas (see Exhibit):

1. Data foundation: Build a rich view of the customer

Many companies have the elements of a relatively complete view of the customer already. But they reside in discrete pockets across the company. Just as a recipe does not come together until all the

Exhibit

Building deeper 1-to-1 relationships with consumers at scale



ingredients are combined, it is only when data is connected that it becomes ready to use. The CDP takes the data a company already has, combines it to create a meaningful customer profile, and makes it accessible across the organization.

"Feeding" the CDP starts by combining as much data as possible and building on it over time. Creating models that cluster customer profiles that behave and create value in similar ways requires advanced analytics to process the data and machine learning to refine it. Over time, as the system "learns," this approach generates ever-more-granular customer subsegments. Signals that the consumer leaves behind (e.g., a site visit, a purchase on an app, interest expressed on social media) can then expand the data set, enabling the company to respond in real time and think of new ways to engage yet again. Furthermore, the insights gleaned extend beyond a customer's response to a specific campaign, for example by driving more targeted product development.

A number of companies we're familiar with, struggling to truly understand their customers who make infrequent purchases, combine their own CRM data with Facebook consumer data to build look-alike models. This helps identify the highest-value prospects most likely to buy in their category. Increased targeting through display ads on and off Facebook can yield 50 to100 percent higher returns than from the average Facebook audience. Mapping third-party data (when it exists) to customer segments via a data-management platform (DMP) can enhance the experience for both known and anonymous digital consumers, leading to improvements in engagement and conversion, measured in net promoter score, acquisition, and lifetime value.

2. Decisioning: Mine the data to act on the signals

The decisioning function enables marketers to decide what is the best content to send to a given customer for a given time and channel. Customers are scored based on their potential value. A set of business rules and regression models (increasingly done through machine learning) then matches specific messages, offers, and experiences to those customer scores, and prioritizes what gets delivered and when. This allows companies to make major improvements in how they engage with their customers by developing more relevant, personalized engagement, within a single channel or across channels, based on a customer's behavioral cues. Those signals can be basic, such as "cart abandoned" or "browsed but didn't buy," or more nuanced, such as activity by segment and time of day, gleaned from mining customer data. In effect, these signals become triggers that invoke an action. A decisioning engine develops a set of triggers and outcomes based on signals and actions the company takes in response.

For example, one multichannel retailer discovered that many consumers made a purchase on the website just once per year. Further analysis revealed those same customers tended to return to browse the site a few days after purchase. The company now takes advantage of this window of opportunity to send tailored, targeted messaging, rather than risk losing the customer for another year. This approach doubled the open rate of its emails—from 10 to15 percent for generic targeted communications to 25 to 35 percent for real-time, "trigger-based" communications acting on consumer signals.

More sophisticated companies build up a decisioning model that works across all distribution channels. That requires advanced modeling and analytics techniques to identify the impact of one channel on another as a customer proceeds along his/her decision journey. A travel company took this approach recently and saw coordinating messages across channels drive a 10 to 20 percent incremental boost in conversion rates and customer lifetime value.

Effective decisioning is based on repeated testing that validates and refines hypotheses and outcomes. Over time, these can become increasingly sophisticated as models and algorithms build on each other. One telecommunications company has been testing different offers to different groups: millennials, customers in specific cities, previous owners of a specific device, groups of relatives, and people who viewed a specific web page in the last three days. As complex as this may seem, a semi-automated decisioning engine prioritizes the offers and experiences proven to have the highest rate of return.

Data-activation self-assessment

This self-assessment can help company leaders develop a benchmark for measuring their progress on their data- activation journey.

Data foundation

How comprehensive is your view of the consumer across all your internal data sets, and how close to real time are those data feeds being updated?
Lagging: We do not use any data for personalization.
Basic: Data-driven personalization is mostly focused on transaction data, and/or anonymous third-party data. Data is manually updated daily or weekly.
Leading: Rich view of consumer across most touchpoints (e.g., transactions, media, clickstream, servicing/care). Data is actively used for personalization. Data is real time or refreshed multiple times per day.

Decisioning

What types of models are you activating across channels? Who manages your models?
Lagging: We are not using any propensity models to enhance targeting or to trigger personalized experiences.
Basic: We have basic propensity models that are used on a limited basis and not used widely in digital. We have limited or no dedicated data-science resources to manage models.
Leading: We have multiple propensity models to predict value creation or destruction for a given customer interaction, and most digital messaging is triggered by these propensity scores. Our models are managed by in-house data-science resources. We currently or soon will use machine learning to further fine-tune models.

Design

How often do you test offers and messages?
Lagging: We do limited tests and do not update our offers frequently.
Basic: Tests are set up and deployed manually. We analyze performance weekly or monthly and optimize periodically.
Leading: We run triggered A/B and/or multivariate tests daily.

Distribution

How are your marketing technology platforms integrated with your data systems?
Lagging: We have not optimized our martech stack and/or rely solely on the platforms our agency manages on our behalf.
Basic: We manually batch upload data to our martech systems, and we are able to deliver personalized experiences to broad customer segments in some channels. Response and transaction data is batch delivered back into the CDP.
Leading: We have API connections between our customer-data platform and our martech systems. All response and transaction data is fed back in a closed loop into our customer data platform.

This allows the telco to scale the results of dozens of tests without fear of inconsistent customer experiences or conflicting offers.

3. Design: Craft the right offers, messages, and experiences at speed

Understanding your customers and how to engage them counts for little without the content to actually deliver to them. Designing great offers, however, is hampered by the fact that functions and departments within companies tend to operate as mini fiefdoms. The owners of each channel test and engage consumers exclusively within their own channel. Real benefits can occur only when companies shift to "war rooms" of people from relevant functions (marketing, digital, legal, merchandising, and IT/DevOps) who focus on specific consumer segments or journeys.

These teams have clear ownership of consumer priorities and responsibility for delivering on them. The cross-functional team continually develops new ideas, designs hypotheses for how to engage customers, devises experiments, and creates offers and assets. Analytics help size opportunities, test impact, and derive insights from tests. That content is then tagged, so that it can be associated with a trigger and be ready to go when needed. Just three months after launching its war room, one large multichannel retailer saw its testing speed go from 15 to 20 weeks to two to three weeks, and testing volume increase from four to six per month to 20 to 30 per month.

4. Distribution: Deliver experiences across platforms

Distribution systems are simply the "pipes" that deliver the ad or content to the end user (e.g., ad server, DSP, or content management platform). Often they can be quite manual and just blast out communications to wide segments of people with little tailoring. But connect the CDP engine, with its predetermined triggers and tagged content, to these distribution systems, and a formerly blunt marketing instrument becomes a far more directed one sending specific messages to distinct customer subsegments across all addressable channels. Sophisticated businesses have developed a library of APIs to help tie the CDP into the "martech stack"—the marketing technologies that deliver and track experiences. Integrating the stack this way creates a feedback loop that sends customer response, engagement, and conversion data back into the CDP.

Implementing the data-activation framework

Not all data-activation efforts are created equal. We recommend using a case-driven approach, maintaining a backlog of tests ranked by opportunity, quantifying the impact of each potential use case, and balancing it with the level of effort required to implement it.

Unlike a wholesale IT transformation, deploying a CDP isn't a replacement of current customer-data systems, but rather an operational solution that can piggyback on existing systems. In our experience, many marketers already have a large part of the marketing-technology equation in house; they're just not using it properly.

The promise of data-activated, one-to-one marketing is not only possible but is now increasingly expected by today's customers. It is now the key to transforming simple customer transactions into enduring relationships.

Julien Boudet is a partner in McKinsey's Seattle office, **Brian Gregg** is a partner in the San Francisco office, **Jason Heller** is a partner in the New York office, and **Caroline Tufft** is a partner in the London office.

The authors wish to thank Sean Flavin, McKinsey expert in personalization, for his contributions to this article.

Operations

Ops 4.0: Fueling the next 20 percent productivity rise with digital analytics

Mercedes Goenaga, Philipp Radtke, Kevin Speicher, and Rafael Westinner

Business needs to raise productivity more than ever. Thanks to innovations in digitization and analytics, four new methodologies can yield the productivity breakthroughs organizations need.

Business is now in the midst of the most significant disruption in decades. This epochal transformation has been driven largely by technological changes—big data and advanced analytics, additive manufacturing, the Internet of Things, robotics, and artificial intelligence—collectively described as the fourth industrial revolution. Arriving at dizzying speed, its consequences are already evident across sectors: competition is intensifying not just within industries but also between them. Think of Apple assembling an autonomous-vehicle business or Tesla moving into power supply. And then there are the aggressive, agile start-ups, with business models that ignore conventional constraints.

Together, these pressures are both intensifying the long-standing imperative to raise productivity (see sidebar "What is productivity?") and leaving much less room for error. Yet they also involve novel tools and methods—for example, vastly increased connectivity and the Internet of Things—with a huge potential for realizing new levels of productivity across the entire value chain.

In 2016, about 17.6 billion devices were connected to the Internet. By 2025, that figure will probably jump to about 80 billion, at a rate of 152,000 a minute.

The difficulty, of course, is to take advantage of these technological breakthroughs in ways that lead to comparable performance breakthroughs. This has never been easy to do. In 1987—more than 30 years after businesses started using mainframes—Nobel Prize–winning economist Robert Solow famously noted, "You can see the computer age everywhere but in the productivity statistics."

Businesses—indeed, societies—cannot afford another 30-year wait for significantly better productivity. They need gains on the order of 20 percent or more, and they need them much sooner. But the problem now, as a generation ago, is that organizations too often overinvest in technology while underinvesting in the human capabilities needed to make it useful.

The real lesson from technology leaders is that they apply it judiciously, as part of a broader transformation of the way they do business, starting with their people. How companies transform themselves depends, to a great extent, on the capabilities they need most.

We see four primary structures, which collectively become Operations 4.0 (Exhibit 1).

- **Product driven.** For organizations whose strategic imperative is to design and launch products more effectively, advanced analytics combines with design to value, becoming analytics to value, or AtV.

- **Journey driven.** Many organizations have already seen a dramatic impact from applying lean management's end-to-end perspective to their customer journeys. Digital technologies and agile processes let organizations make these changes more easily, quickly, and sustainably—and on a greater scale, with a bigger impact—than ever before. Together, the technologies and processes form EdgE, or end-to-end digital enablement.

- **Cost and budget driven.** Traditional cost-control measures have often been a blunt instrument at best, but a more sophisticated analysis required too much data and coordination to be practical. Now, sophisticated analytics techniques make zero-based budgeting, or ZBB, more feasible, flexible, and profitable than ever.

- **Manufacturing driven.** To help companies reach new levels of resource productivity and effectiveness, digital manufacturing connects novel and existing data sources with smarter machines and new process technologies.

Exhibit 1

Operations 4.0 encompasses four approaches to achieve productivity breakthroughs.

Cost and budget driven

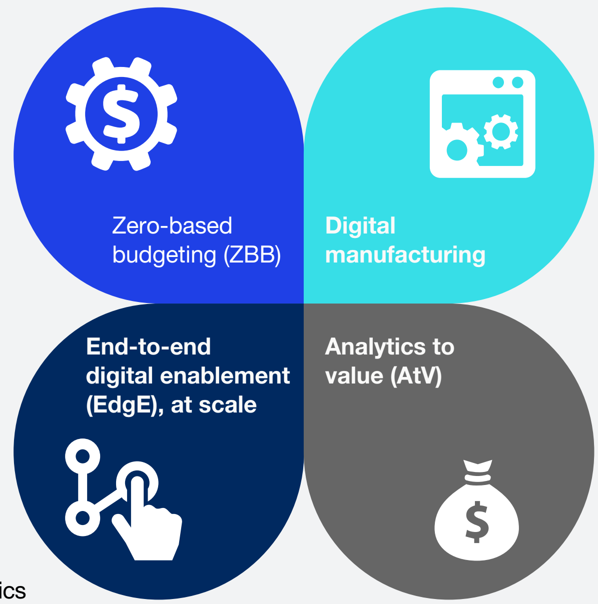
ZBB data aggregation and analysis

Smart capital-expense optimization

Journey driven

End-to-end digitization of processes

Intelligent process automation/robotics



Enabled by big data, digital, and analytics

Manufacturing driven

Big data–enabled predictive maintenance

Yield, energy, and throughput analytics

Experiential learning for digital manufacturing

Real-time performance management

Digital analytics diagnostics

Supply chain 4.0

Product driven

Complexity management with big data

Digital analytics-enabled procurement

Advanced cleansheets

Design-based product optimization

R&D productivity through machine learning

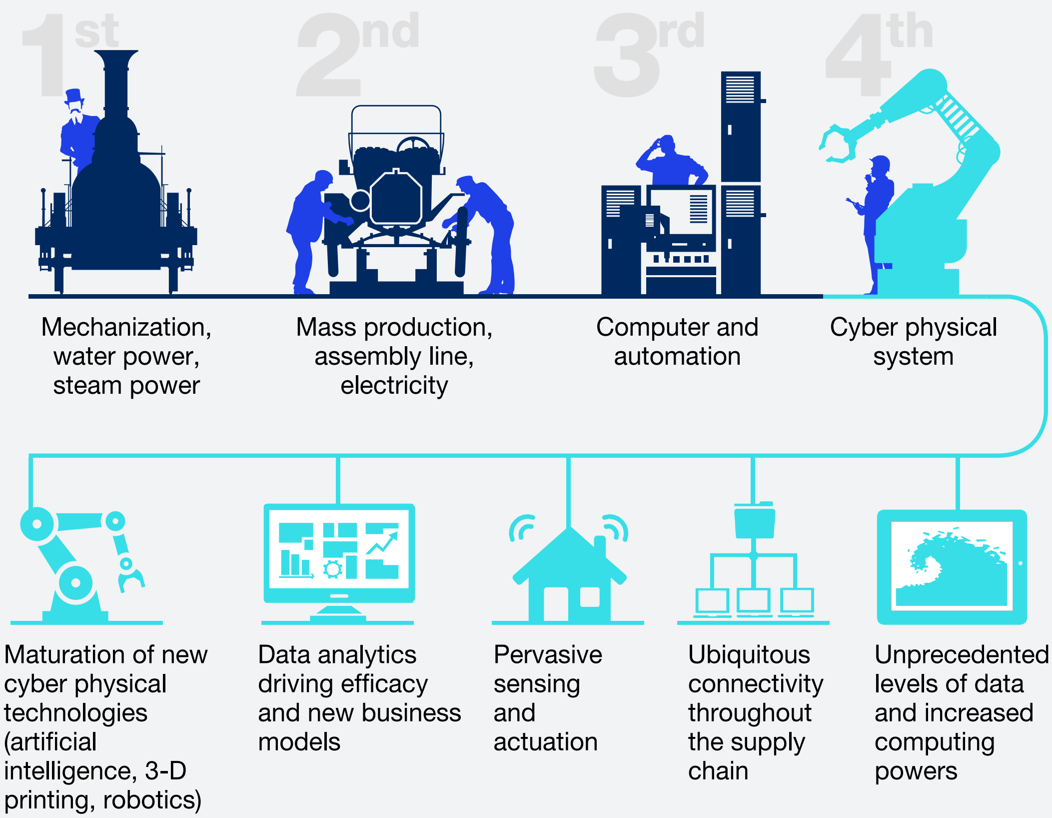
How digital analytics fuels the next 20 percent productivity rise

The fourth industrial revolution's digital analytics can support a productivity leap because it generates so many distinct opportunities (Exhibit 2). Among the simplest is faster acceleration: changes happen more quickly and organizations can do more things in less time. Higher efficiency means that these changes require fewer resources, while enhanced effectiveness gives the changes greater effect. Increased predictability—achieved, for example, through more accurate forecasting based on unstructured data—lets organizations plan their moves more consistently and respond with greater agility. Finally, deeper engagement at every level yields denser, larger resource networks, which reinforce new behaviors and help build a transformation's scale.

These areas of impact all combine in different ways, depending on an organization's starting point and the type of transformation it undertakes. Together, they make it more likely that the changes will keep performance improving year after year.

Exhibit 2

In the fourth industrial revolution, digital analytics enables a new level of operational productivity.



Source: *Forbes*; World Economic Forum

Product driven: From cost to design to analytics-to-value

Now that essential product functions have become increasingly commoditized, product design has emerged as a crucial source of differentiation. But the best companies have already extracted many obvious sources of advantage from this. The next level of product optimization therefore not only combines the latest design thinking with multiple sources of data but also exploits sophisticated advanced-analytics methodologies to generate insights about potential cost and value improvements. For example, computer-aided design tools linked to vast pools of procurement data, social-media activity, and cost and complexity benchmarks can allow a company to quickly identify designs that maximize profitability while minimizing wasted time and effort.

Such breakthroughs are not just for the consumer sector. One of the world's largest industrial conglomerates brings these ideas to life with products meant not for individuals but for utilities—whose traditional business model has been upended by renewable (and increasingly customer-generated) energy sources and more sophisticated consumers. The conglomerate's improvement target: within four years, cut delivery lead times by more than half, defend and increase market share, and raise profit margins by about 30 percent.

Complexity management. In utilities, as in much of today's business world, decades of acquisitions have left many companies managing dozens of systems—especially IT systems—that never get fully integrated. Meanwhile, product proliferation is a constant battle as small variants in specifications generate hundreds of mostly overlapping SKUs. Standard methodologies for combating this complexity not only take vast amounts of time and effort but also may not even identify the right changes. Yet with new digital analytics tools, the conglomerate completed an analysis, in just two weeks instead of several months, that identified specific commonalities the company could use to reduce variations among product families, subsystems, and components.

Analytics and automation. Analytics has made procurement a much more promising target for savings by tapping a previously impractical data source: the procurement and engineering departments' own bills of materials. New tools can upload thousands of records, held around the world in dozens of local languages and part-numbering structures, to find potential commonalities and opportunities to negotiate better pricing. An early step toward artificial intelligence, robotic process automation, can then allow software "robots" to take over tedious processes, such as collating information from disparate systems for complex forms, and thus frees people to focus on work that uses their judgment and experience. Finally, combining multiple data streams—such as on actual spending, product cost structures, sales, and so forth—into a data "lake" allows sophisticated algorithms to engage in optimization dynamically, enabling constant adjustments as conditions change.

Powerful portfolio analysis. Together, techniques such as these can generate a much more detailed analysis of an entire product portfolio. With slight modifications, the conglomerate found it could eliminate 15 to 80 percent of product variants within a category. Already, costs have improved by approximately 30 percent.

Journey driven: Minimizing the middle

First in manufacturing, and later in virtually every sector from banking to government to warehousing, the disciplines collectively known as lean management have enabled organizations to focus ever more tightly on doing only whatever creates value that customers are willing to pay for. New technologies are making these disciplines more critical and powerful than ever. An insurer, for example, reduced its time to market for new products from 18 months to three, and

What is productivity?

Productivity can be a slippery concept. Of the several definitions the *Webster's Third New International Dictionary Unabridged* (Merriam-Webster) gives, the one we hear most often in an operations context is "the degree of effectiveness of industrial management in utilizing the facilities for production; especially: the effectiveness in utilizing labor and equipment." This is admirably clear but enforces a narrow perspective on what productivity really implies.

We believe that a broader alternative—"output per unit of productive effort"—gives a better idea of what productivity can encompass. Everything a company does is productive effort that can be measured against output. Productivity therefore includes every aspect of operational excellence, from the generation of an idea for a product through its manufacture, sale, maintenance, and, potentially, dismantling and recycling once it becomes obsolete.

a government body replaced 50 legacy platforms with a new enterprise-resource-planning system delivered on time and on budget.

Understanding the complete journey. First, technologies are making it easier and faster for organizations of all kinds to see their processes as their customers (or constituents) do—not as a series of departments, but as journeys that have a start, a middle, and an end. The trouble is usually in the middle, as customers struggle with redundant steps, poor communication, and delays because a company's functions don't coordinate their activities.

Analytics tools let organizations see exactly how customers move from one point to another, both within and between channels. Those insights can help a company fix really basic issues—such as simplifying online registration forms or reminding customers to bring a government ID when they pick up products in person—that make a big difference in the customer experience.

Speeding the journey with digital. Next, the organization starts thinking about where technology can change processes more fundamentally. The goal is to ensure that digitization fits in with the way today's customers actually behave rather than the way companies might have built processes in an analog world. Links among back-office systems, for example, help prepopulate forms for existing customers—or eliminate the forms entirely because the information is already available and robotic process automation has already brought all of it together.

Building agility for the future. Last year's breakthroughs quickly become this year's table stakes. To keep up, organizations must behave as tech leaders do, by learning how to refine a moving target: ruthlessly trimming ideas to their "minimum viable product" core, testing and improving them, and then adding new features in the same cycle.

One global financial institution has already transformed customer journeys covering about 80 percent of its interactions with clients. Its culture is changing, as well. Faster decision making and a higher level of comfort with the build-test-revise cycle is helping improve customer experience, while increasing sales conversions in highly competitive product categories by between 4 to 8 percent.

Cost and budget driven: Zero as hero

First developed almost half a century ago, zero-based budgeting has already proved its power by achieving billions of dollars in lasting cost reductions. Several of the most high-profile examples involve companies in the consumer sector, where slow growth called for drastic action. But in other sectors, skepticism remains common; executives fear that these ideas will take up so much time and attention that the organization will be left doing little else, and that as a result, ZBB will become a one-time exercise.

Yet today, some companies are already putting simple ZBB-based digital analytics tools in the hands of virtually every employee with budgetary responsibility. In this way, they are building speed, scale, and sustainability throughout the ZBB process and realizing cost savings of 10 to 25 percent in the first year, with additional savings thereafter.

Making better trade-offs. The proof is in the budgeting. Historically, applying ZBB meant developing detailed templates and collating hundreds of spreadsheets with different data structures and different levels of quality and granularity. Today's integrated planning platforms build in the required data—including, in some cases, detailed benchmarks stretching back years or even decades. Managers can readily make complex trade-offs that balance policy considerations (say, a preference for nonrefundable tickets) with variables such as fares, average lodging costs, time of year, country, and traveler seniority. The integrated platform makes it easy to iterate individual budget-plan components in any direction: top down (from executives to planners) and vice versa, in a tight cycle. The result is a budget rooted in detailed insights, with clear personal responsibility for each item—an immense cultural change.

Sustaining a new culture. These improvements can easily keep going year after year. One global company built a center of excellence where ZBB analysts work with cost-category owners to update categories and prices so that each year's budget update uses the latest optimized data. That human investment helps ensure that the company's continuous improvement extends to ZBB itself.

Manufacturing driven: More value with less waste

The final major set of opportunities—how companies manufacture goods—arises from the shop floor. Much of the business world's attention has focused on new manufacturing technologies, such as 3-D printing. What promises to have an even greater impact is the way these innovations combine with less dramatic but equally far-reaching developments, such as the emergence of cheap Internet-linked sensors (a highly pragmatic application of the Internet of Things) and user-friendly advanced-analysis tools. Together, these technologies, which give human beings an unprecedented degree of understanding and control over forbiddingly complex processes, have an enormous economic effect.

Finding opportunity—fast. One large high-tech manufacturer illustrates the potential of combining these tactics. Facing heightened competition and eroding margins, the leaders of the company knew that it needed the improvements promised by digital technologies and advanced analytics. The first step was a 48-hour diagnostic: specialists gathered data on the company's most important production equipment, revealing many gaps in basic manufacturing hygiene. Equipment downtime was unacceptably high, production quality uneven, and overall efficiency much lower than what competitors had achieved. Until the company addressed these issues, adding new technologies would be a waste.

Upgrading the supply chain. Within operations, demand planning and supply-chain logistics have long been at the forefront in applying digital technologies. Now the bar is rising still higher: customers increasingly expect the quality and service breakthroughs that Operations 4.0 technologies make possible. No-touch order processing and real-time, reliable replanning, for example, enable a better customer experience. But they also mean erasing the traditional boundaries between the supply chain, manufacturing, and fulfillment, as 3-D printing reconfigures logistics and advanced robotics support smart warehouses.

Changing people first. But the root causes of the challenges centered not on equipment but on people—especially managing performance. A new digital system now does so all the way from the factory floor to the CEO level, allowing everyone to see and fix gaps at all times. The company quickly made the equipment about 20 percent more effective, with corresponding increases in quality. Most important, it could then start restructuring about 50 percent of its manufacturing processes to enable technologies such as data-driven predictive maintenance (reducing downtime by an additional 30 percent), a digitized quality system, advanced process controls, robotic in-plant logistics, and automation, as well as human–machine collaboration (Exhibit 3). Throughout the initiative, human capital has been preserved, subject only to natural attrition and redeployment of people to other areas of the company.

▼▼▼

Combining people, novel digital technologies, and advanced analytics can yield a new breakthrough in productivity if companies learn to weave them all together. Doing so will require

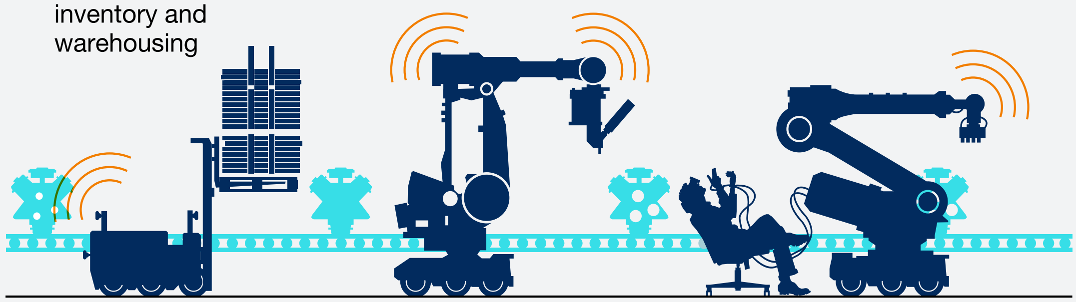
Exhibit 3

The factory of the future combines technologies that are available today.

1. Automated in-plant logistics handle inventory and warehousing

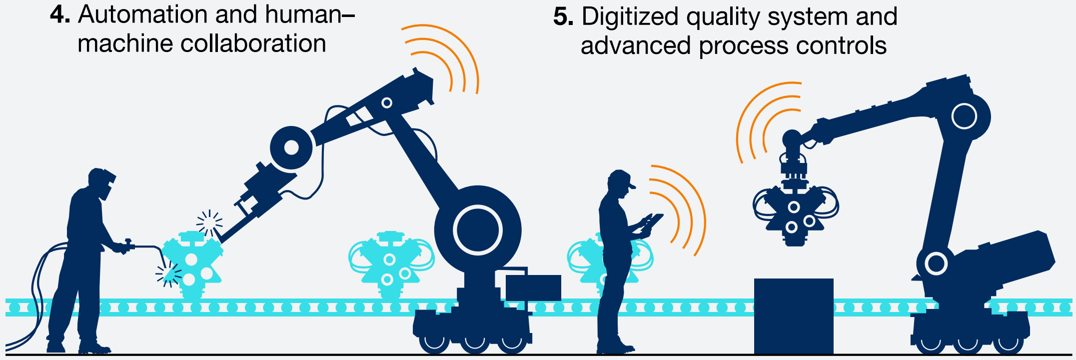
2. Data collection across the supply chain

3. Data-driven predictive maintenance



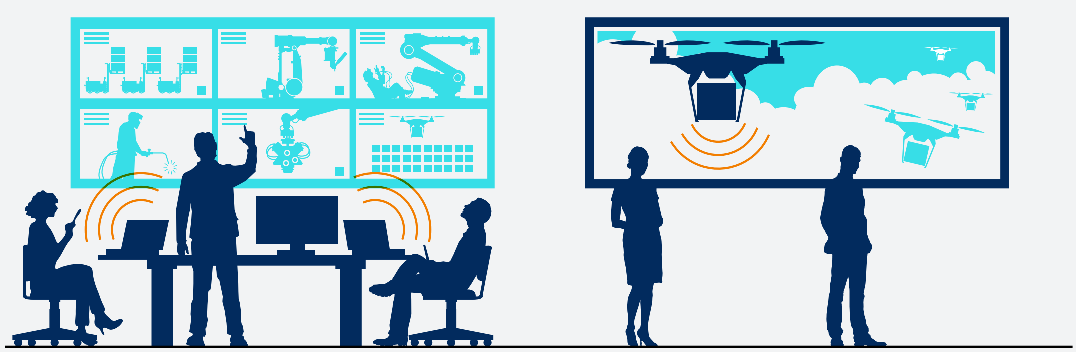
4. Automation and human–machine collaboration

5. Digitized quality system and advanced process controls



6. Digital performance-management system working with enablers and IT infrastructure

7. Smart planning and agile operations deliver products directly to customer using latest technologies and tracking



Source: McKinsey Global Institute analysis sustained commitment across the entire organization: to coordinate stakeholders with diverging agendas (IT leaders struggling with legacy systems, business heads controlling "their" data), to help people change their mind-sets (from intuition to reasoning, or from easy generalities to hard specifics), and to create entirely new capabilities (in extracting insights from data, and crafting actions from insights). But with the right support, the power of Operations 4.0 becomes far too great for leaders to let it pass by.

Mercedes Goenaga is a specialist in McKinsey's Paris office, **Philipp Radtke** is a senior partner in the Munich office, **Kevin Speicher** is a senior partner in the Chicago office, and **Rafael Westinner** is a partner in the Madrid office.

The authors wish to thank Joao Dias, Manuela Mareva-Forsell, Yogesh Malik, Abhishek Shirali, Jan Wüllenweber, and David Zhu for their contributions to this article.

Organization

Using people analytics to drive business performance: A case study

Carla Arellano, Alexander DiLeonardo, and Ignacio Felix

A quick-service restaurant chain with thousands of outlets around the world is using data to drive a successful turnaround, increase customer satisfaction, and grow revenues.

People analytics—the application of advanced analytics and large data sets to talent management—is going mainstream. Five years ago, it was the provenance of a few leading companies, such as Google (whose former senior vice president of people operations wrote a book about it). Now a growing number of businesses are applying analytics to processes such as recruiting and retention, uncovering surprising sources of talent and counterintuitive insights about what drives employee performance.

Much of the work to date has focused on specialized talent (a natural by-product of the types of companies that pioneered people analytics) and on individual HR processes. That makes the recent experience of a global quick-service restaurant chain instructive. The company focused the power of people analytics on its frontline staff—with an eye toward improving overall business performance—and achieved dramatic improvements in customer satisfaction, service performance, and overall business results, including a 5 percent increase in group sales in its pilot market. Here is its story.

The challenge: Collecting data to map the talent value chain

The company had already exhausted most traditional strategic options and was looking for new opportunities to improve the customer experience. Operating a mix of franchised outlets, as well as corporate-owned restaurants, the company was suffering from annual employee turnover significantly above that of its peers. Business leaders believed closing this turnover gap could be a key to improving the customer experience and increasing revenues, and that their best chance at boosting retention lay in understanding their people better. The starting point was to define the goals for the effort and then translate the full range of frontline employee behavior and experience into data that the company could model against actual outcomes.

Define what matters. Agreeing in advance on the outcomes that matter is a critical step in any people-analytics project—one that's often overlooked and can involve a significant investment of time. In this case, it required rigorous data exploration and discussion among senior leaders to align on three target metrics: revenue growth per store, average customer satisfaction, and average speed of service (the last two measured by shift to ensure that the people driving those results were tracked). This exercise highlighted a few performance metrics that worked together and others that "pulled" in opposite directions in certain contexts.

Fill data gaps. Internal sources provided some relevant data, and it was possible to derive other variables, such as commute distance. The company needed to supplement its existing data, however, notably in three areas (Exhibit 1):

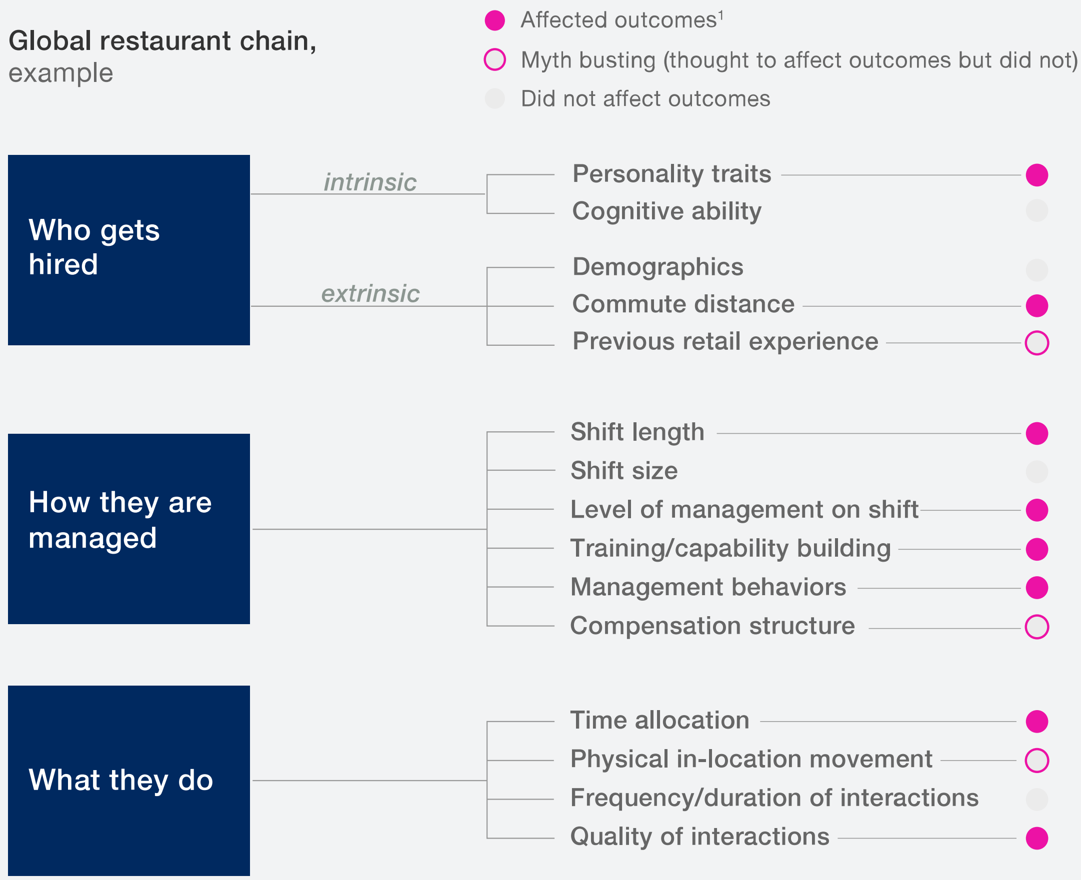
- First was selection and onboarding ("who gets hired and what their traits are"). There was little data on personality traits, which some leaders thought might be a significant factor in explaining differences in the performance of the various outlets and shifts. In association with a specialist in psychometric assessments, the company ran a series of online games allowing data scientists to build a picture of individual employees' personalities and cognitive skills.

- Second was day-to-day management ("how we manage our people and their environment"). Measuring management quality is never easy, and the company did not have a culture or engagement survey. To provide insight into management practices, the company deployed McKinsey's Organizational Health Index (OHI), an instrument through which we've pinpointed 37 management practices that contribute most to organizational health and long-term performance. With the OHI, the company sought improved understanding of such practices and the impact that leadership actions were having on the front line.

- Third was behavior and interactions ("what employees do in the restaurants"). Employee behavior and collaboration was monitored over time by sensors that tracked the intensity of physical interactions among colleagues. The sensors captured the extent to which employees

Exhibit 1

Analysis identified which employee features correlated to the desired outcomes.



[1]Targeted outcomes were customer-satisfaction scores by shift, revenue growth by store, and speed of service by shift.

physically moved around the restaurant, the tone of their conversations, and the amount of time spent talking versus listening to colleagues and customers.

The insights: Challenging conventional wisdom

Armed with these new and existing data sources—six in all, beyond the traditional HR profile, and comprising more than 10,000 data points spanning individuals, shifts, and restaurants across four US markets, and including the financial and operational performance of each outlet—the company set out to find which variables corresponded most closely to store success. It used the data to build a series of logistic-regression and unsupervised-learning models that could help determine the relationship between drivers and desired outcomes (customer satisfaction and speed of service by shift, and revenue growth by store).

Then it began testing more than 100 hypotheses, many of which had been strongly championed by senior managers based on their observations and instincts from years of experience. This part of the exercise proved to be especially powerful, confronting senior individuals with evidence that in some cases contradicted deeply held and often conflicting instincts about what drives success. Four insights emerged from the analysis that have begun informing how the company manages its people day to day.

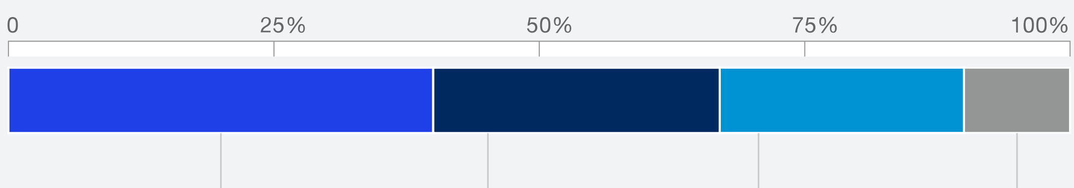
Personality counts. In the retail business at least, certain personality traits have higher impact on desired outcomes. Through the analysis, the company identified four clusters or archetypes of frontline employees who were working each day: one group, "potential leaders," exhibited many characteristics similar to store managers; another group, "socializers," were friendly and had high emotional intelligence; and there were two different groups of "taskmasters," who focused on job execution (Exhibit 2). Counterintuitively, though, the hypothesis that socializers—and hiring for friendliness—would maximize performance was not supported by the data. There was a closer correlation between performance and the ability of employees to focus on their work and minimize distractions, in essence getting things done.

Careers are key. The company found that variable compensation, a lever the organization used frequently to motivate store managers and employees, had been largely ineffective: the data

Exhibit 2

Frontline employees fell into four personality archetypes.

Distribution of employees at a global restaurant chain



Socializers	Potential leaders	Entrepreneurial taskmasters	Conservative taskmasters
High EQ,[1] more altruistic and trusting	High EQ	Lower EQ, less altruistic	Low EQ
Risk takers, highly spontaneous	Good at multitasking	Good at planning and executing	Good at planning and execution
	Follow up with others	Higher appetite for risk and innovation	Very focused, not good at multitasking
	Most like high-performing general managers		Conduct work within boundaries provided (not risk seeking)

[1]Emotional Quotient, a measure of self-awareness and sensitivity to others.

suggested that higher and more frequent variable financial incentives (awards that were material to the company but not significant at the individual level) were not strongly correlated with stronger store or individual performance. Conversely, career development and cultural norms had a stronger impact on outcomes.

Management is a contact sport. One group of executives had been convinced that managerial tenure was a key variable, yet the data did not show that. There was no correlation to length of service or personality type. This insight encouraged the company to identify more precisely what its "good" store managers were doing, after which it was able to train their assistants and other local leaders to act and behave in the same way (through, for example, empowering and inspiring staff, recognizing achievement, and creating a stronger team environment).

Shifts differ. Performance was markedly weaker during shifts of eight to ten hours. Such shifts were inconsistent both with demand patterns and with the stamina of employees, whose energy fell significantly after six hours at work. Longer shifts, it seems, had become the norm in many restaurants to ease commutes and simplify scheduling (fewer days of work in the week, with more hours of work each day). Analysis of the data demonstrated to managers that while this policy simplified managerial responsibilities, it was actually hurting productivity.

▼▼▼

The results (so far)

Four months into a pilot in the first market in which the findings are being implemented, the results are encouraging. Customer satisfaction scores have increased by more than 100 percent, speed of service (as measured by the time between order and transaction completion) has improved by 30 seconds, attrition of new joiners has decreased substantially, and sales are up by 5 percent.

We'd caution, of course, against concluding that instinct has no role to play in the recruiting, development, management, and retention of employees—or in identifying the combination of people skills that drives great performance. Still, results like these, in an industry like retail—which in the United States alone employs more than 16 million people and, depending on the year and season, may hire three-quarters of a million seasonal employees—point to much broader potential for people analytics. It appears that executives who can complement experience-based wisdom with analytically driven insight stand a much better chance of linking their talent efforts to business value.

Carla Arellano is a vice president of, and **Alexander DiLeonardo** is a senior expert at, People Analytics, a McKinsey Solution—both are based in McKinsey's New York office; **Ignacio Felix** is a partner in the Miami office.

The authors wish to thank Val Rastorguev, Dan Martin, and Ryan Smith for their contributions to this article.

Risk

Risk analytics enters its prime

Rajdeep Dash, Andreas Kremer, Luis Nario, and Derek Waldron

All the ingredients are in place for unprecedented advances in risk analytics. Now it's up to banks to capture the opportunities.

With the rise of computing power and new analytical techniques, banks can now extract deeper and more valuable insights from their ever-growing mountains of data. And they can do it quickly, as many key processes are now automated (and many more soon will be). For risk departments, which have been using data analytics for decades, these trends present unique opportunities to better identify, measure, and mitigate risk. Critically, they can leverage their vast expertise in data and analytics to help leaders shape the strategic agenda of the bank.

Banks that are leading the analytical charge are exploiting both internal and external data. Within their walls, these banks are integrating more of their data, such as transactional and behavioral data from multiple sources, recognizing their high value. They are also looking externally, where they routinely go beyond conventional structured information, such as credit-bureau reports and market information, to evaluate risks. They query unconventional sources of data (such as government statistics, customer data from utilities and supermarket loyalty cards, and geospatial data) and even new unstructured sources (such as chat and voice transcripts, customer rating websites, and social media). Furthermore, they are getting strong results by combining internal and external data sets in unique ways, such as by overlaying externally sourced map data on the bank's transaction information to create a map of product usage by geography. Perhaps surprisingly, some banks in emerging markets are pioneering this work. This is possible because these banks are often building their risk database from scratch and sometimes have more regulatory latitude.

The recent dramatic increases in computing power have allowed banks to deploy advanced analytical techniques at an industrial scale. Machine-learning techniques, such as deep learning, random forest, and XGBoost, are now common at top risk-analytics departments. The new tools radically improve banks' decision models. And techniques such as natural-language processing and geospatial analysis expand the database from which banks can derive insights.

These advances have allowed banks to automate more steps within currently manual processes—such as data capture and cleaning. With automation, straight-through processing of most transactions becomes possible, as well as the creation of reports in near real time. This means that risk teams can increasingly measure and mitigate risk more accurately and faster.

The benefits—and challenges—of risk analytics

Banks that are fully exploiting these shifts are experiencing a "golden age" of risk analytics, capturing benefits in the accuracy and reach of their credit-risk models and in entirely new business models. They are seeing radical improvement in their credit-risk models, resulting in higher profitability. For example, Gini coefficients of 0.75 or more in default prediction models are now possible.[1] Exhibit 1 lays out the value that analytics can bring to these models.

Some banks are expanding their risk models to new realms. A few have been able to automate the lending process end to end for their retail and SME segments. These banks have added new analytical tools to credit processes, including calculators for affordability or preapproval limits. With this kind of straight-through processing banks can approve up to 90 percent of consumer loans in seconds, generating efficiencies of 50 percent and revenue increases of 5 to 10 percent. Recognizing the value in fast and accurate decisions, some banks are experimenting with using risk models in other areas as well. For example, one European bank overlaid its risk models on its marketing models to obtain a risk-profitability view of each customer. The bank thereby improved the return on prospecting for new revenue sources (and on current customers, too).

A few financial institutions at the leading edge are using risk analytics to fundamentally rethink their business model, expanding their portfolio and creating new ways of serving their customers. Santander UK and Scotiabank have each teamed up with Kabbage, which, using its own partnership with Celtic Bank, has enabled these banks to provide automated underwriting of small-business loans in the United Kingdom, Canada, and Mexico, using cleaner and broader data sets. Another leading bank has used its mortgage-risk model to provide a platform for real estate agents and others providing home-buying services.

1 Gini coefficients measure variation or randomness in a set of values, where 0 is completely random and 1 is perfectly ordered. In a model that predicts default, a Gini coefficient of 0 would indicate that the model is no better than a coin toss, and 1 would indicate that the model's output perfectly predicted the eventual defaults.

Exhibit 1

Analytically enhanced credit models can improve banks' returns in four ways.

Higher interest income from loan business

- Increase loan volume through sales campaigns, with lower turndown rate due to better customer preselection and cross-selling
- Increase margin and loan volume by gradually introducing risk-differentiated offers (eg, packages or prices) and cross-selling of higher-margin products

Lower sales and operating costs

- Targeted and effective origination process (eg, risk prescreening, policy prefilters)
- More efficient underwriting process (eg, digital channels, risk-based differentiated process across products)

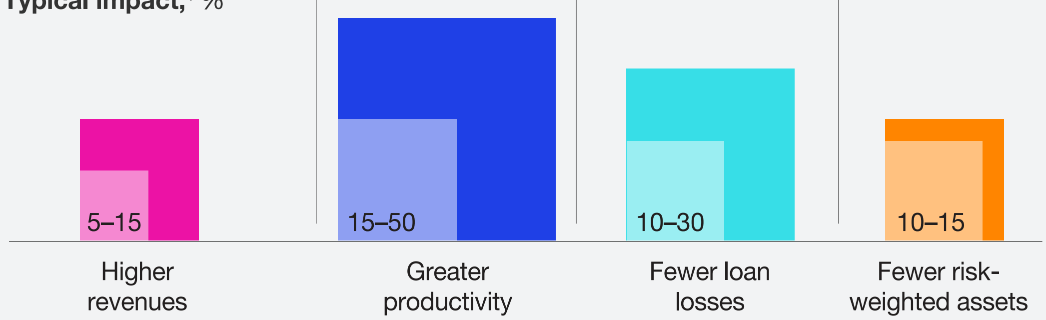
Reduction of relative risk costs

- Better selection of risks, (eg, with combined risk scores, risk clustering of customer segments)
- Improved risk monitoring and early warning across product categories

Improved capital efficiency

- Better calibration and refinement of the models, leading to reduced risk-weighted assets
- Better data cleanliness to accurately represent risk measures and mitigants

Typical impact,[1] %



[1]Impact not additive and depends on the bank's portfolio.

Realizing the potential

For many banks the advantages of risk analytics remain but a promise. They see out-of-date technology, data that is difficult to clean, skill gaps, organizational problems, and unrelenting regulatory demands. The barriers seem insurmountable. Yet banks can get things moving with some deliberate actions (Exhibit 2).

Perhaps the most salient issue is that risk analytics is not yet on the strategic agenda. Bank leaders often don't understand what is really at stake with risk analytics—at times because the analytics managers present highly complex solutions with a business case attached as an afterthought.

Lagging banks miss out on the benefits, obviously, and also put other programs and activities at risk. Initiatives to grow revenue and optimize pricing can founder if imprecise risk assessment of customer segments leads to poor choices. In lending, when risk models underperform, banks often add business

Exhibit 2

Several factors keeping banks from realizing the potential promise of risk analytics should be reexamined.

	Perceived barrier	A better way to think about it
Strategic agenda	• Risk analytics is disconnected from business strategy and often seen as only a technology or regulatory-compliance initiative	• Risk analytics is at the heart of many strategic topics (eg, digital, capital productivity, loan-book health, market entry)
Data and technology	• Unclean, unmatched data means waiting for that never-ending, “nearly complete” data transformation	• The data available can generate high value, often in combination with external data
	• Technological landscape is so complex that a simplification and upgrade is required before doing anything	• The “art of the possible” can produce high-value projects
Skills and organization	• IT group doesn’t have the authority to enforce data-management policies	• The business can take responsibility for data quality, integrity, and access, supported by a strong IT organization
	• Building analytics means hiring scarce, expensive data engineers and scientists	• Banks can move quickly through inorganic growth and partnerships
Regulations	• Regulatory burden does not allow us to focus on anything else, including analytics	• Analytics business cases can tease out surprising synergies between regulatory needs and business aspirations
	• Regulators would not agree with use of advanced models and more advanced data	• Sophisticated, value-generating models can be built even within constraints established by the Basel Committee and the European Union
Change	• Building a model is relatively easy and can be done any time	• Digital economy has “winner takes all” economics; first movers have a huge advantage

rules and policies as well as other manual interventions. But that inevitably degrades the customer experience, and it creates an opening for fintechs to capture market share through a better experience and more precise targeting. Taken to its logical conclusion, it is conceivable that banks might be relegated to "dumb pipes" that provide only financing.

Some nimble risk groups are finding ways through these problems, however. Our analysis suggests these teams have six common behaviors:

1. **Take it from the top,** lifting risk analytics to the strategic agenda. For example, 4 out of 10 strategic actions that HSBC Bank laid out in 2015 rely heavily on risk analytics.

2. **Think big and apply analytics** to every material decision. Capital One is well-known for applying analytics to every decision that it makes, even when hiring data scientists.

3. **Go with what you have.** If data is messy or incomplete, don't wait for a better version or for a "data-lake nirvana." Use the data you have, or find a way to complement it. When Banco Bilbao Vizcaya Argentaria (BBVA) wanted to lend to some clients but lacked information, it partnered with Destacame, a utility-data start-up, to provide data sufficient to support a way to underwrite the customers.

4. **Accumulate skills quickly,** through either rapid hiring or acquisitions and partnerships. Then retain your talent by motivating people with financial and nonfinancial incentives, such as compelling projects. Banks such as BBVA, HSBC, Santander, and Sberbank have launched funds of $100 million and more to acquire and partner with fintechs to add their market share, sophisticated technologies, and people.

5. **Fail often to succeed, iterating quickly** through a series of "minimum viable products" (MVPs) while also breaking down traditional organizational silos. One bank building a fully digital lending product went through six MVPs in just 16 weeks to get to a product it could roll out more broadly.

6. **Use model validation to drive relentless improvement.** Validation teams can be the source of many improvements to risk models, while preserving their independence. The key is for teams to style themselves as the guardian of model performance, rather than the traditional activity of merely examining models.

If banks can master these elements, significant impact awaits. Risk analytics is not the entire answer. But as leading banks are discovering, it is worthwhile in itself, and it is also at the heart of many successful transformations, such as digital risk and the digitization of key processes such as credit underwriting.

Risk-analytics leaders are creating analytic algorithms to support rapid and more accurate decision making to power risk transformations throughout the bank. The results have been impressive. An improvement in the Gini coefficient of one percentage point in a default prediction model can save a typical bank $10 million annually for every $1 billion in underwritten loans.[2] Accurate data capture and well-calibrated models have helped a global bank reduce risk-weighted assets by about $100 billion, leading to the release of billions in capital reserves that could be redeployed in the bank's growth businesses.

Leveraging the six successful behaviors

Nothing succeeds like success. The behaviors we have observed in successful risk-analytics groups provide the guidance.

Take it from the top

Stress testing and regulatory oversight following the 2008 financial crisis have vaulted risk management to the top of the management agenda. Nine years later, and after significant investment, most big banks have regained a handle on their risks and control of their regulatory relations. However, leading banks, recognizing the value from risk analytics, are keeping these programs at the top of their strategic plans, and top leaders are taking responsibility.

Top management attention ensures commitment of sufficient resources and removal of any roadblocks—especially organizational silos, and the disconnected data sets that accompany these divides. Leaders can also keep teams focused on the value of high-priority use cases and encourage the use of cross-functional expertise and cross-pollination of advanced analytical techniques. Good ideas for applications arise at the front line, as people recognize changing customer needs and patterns, so banks must also build and maintain lines of communication.

Think big and apply analytics

For some time, analytics has played an important role in many parts of the bank, including risk, where a host of models—such as the PD, LGD, and EAD[3] models used in the internal ratings-based approach to credit risk—are in constant use. What's new is that the range of useful algorithms has greatly expanded, opening up dozens of new applications in the bank. Many small improvements to material decisions can really add up. An obvious example is algorithmic trading, which has transformed several businesses. Already by 2009, for example, it accounted for 73 percent of traded volume in cash equities. An expansion of automated credit decisions and monitoring has allowed banks to radically improve customer experience in residential mortgages and other areas. Banks in North and South America are using advanced-analytics models to predict the behavior of past-due borrowers and pair them with the most productive collections intervention.

These and other important examples are shown in Exhibit 3. What's important is that leading banks are putting analytics to work at every step of these and many other processes. Any time a decision needs to be made, these banks call on risk analytics to provide better answers. Even as they expand the applications of risk analytics, however, leading banks also recognize that they need to strengthen their model risk management to deal with inherent uncertainties within risk-analytics models, as these make up the largest share of risk-related decisions within banks.

---

2 Assuming a base Gini coefficient of 0.50 and an observed default rate of 5 percent.

3 Probability of default, loss given default, and exposure at default.

Go with what you have

Messy, repetitive, and incomplete databases are a reality—but need not be an excuse. Rather than waiting for improvements in the quality, availability, and consistency of the bank's systems and the data they produce, leading risk-analytics teams ask what can be done now. This might involve using readily available data in the bank to immediately build a core analytic module, onto which new modules are integrated as new data sources become available. Alternatively, integrating two

Exhibit 3

Rapid innovation in eight use cases is powered by advanced analytics.

Credit risk	Description	Use cases
1	Underwriting	**Make better underwriting decisions** by using deep-learning algorithms to process vast amounts of data and more accurately quantify the risk of default
2	Credit-line management	**Reduce charge-off losses** by offering an optimal line to each client that is determined by machine-learning algorithms using the latest information about the client (eg, credit score) and local market (eg, home values)
3	Collections	**Increase recoveries** by making the right offer, at the right time, and through the right channel, with a recommendation engine and decision flow powered by four machine-learning algorithms
Operations risk		
4	Payment-fraud detection	**Identify and review high-risk payments before they are executed** by using input from fraud investigators to tune powerful machine-learning algorithms that pinpoint the highest-risk transactions
5	Anti–money laundering	**Quickly suspend money-laundering** operations using a longitudinal view of payment pathways to identify the patterns most indicative of money laundering, and accelerate reviews with powerful investigative tools
Trading risk		
6	Contract compliance	Automate the **extraction and storage of data** from millions of trading contracts for regulatory compliance using leading-edge image-recognition and machine-learning algorithms
7	Trade surveillance	**Identify high-risk traders** by monitoring their behavior with sophisticated natural language-processing algorithms that recognize themes in trader communications that are markers of conduct risk
Model risk		
8	Model validation	Apply **rigorous and efficient model-validation processes** for traditional and advanced models that meet regulatory expectations and adhere to industry benchmarks for model risk management

or more of the data sets on hand can generate significant value. These approaches hasten new analytical models to market, while at the same time helping the bank gather information as it forms a credit relationship with customers.

Furthermore, leading banks supplement their resources with external data—once they have established that this offers clear additional value. Some US fintechs, for example, obtain customer permission to comb financial data and create a sanitized database that banks can use to make accurate risk decisions based on cash-flow patterns. A bank in Central America built a credit-approval system for unbanked customers based on data collected from supermarket loyalty cards. The bank used data such as frequency of shopping and the amount that customers typically spent per visit to estimate customers' ability to repay debt. Even better for banks, many external data are free. In some markets, micromarket information such as house prices by postal code or employment by district is available, and can be mined for insights into creditworthiness of customers, especially small businesses. Conducting geospatial analytics on this information can also provide valuable insights (for example, proximity to a coffee-chain outlet would reveal foot traffic for a retail shop). Banks have also started analyzing unstructured data sets, such as news articles, feedback sites, and even social-network data.

Leading banks apply two tests before acquiring external data: Will it add value, typically through combination with other data sets? And does it conform with the bank's regulatory and risk policies? Consumer-protection regulations restrict the type of data that banks can use for risk-analytics applications, such as lending and product design.

While the practices outlined here will yield fast impact from messy, repetitive, and incomplete databases, most banks would still benefit from establishing sound data governance in parallel (and sometimes are required to do so under data regulations such as BCBS 239).

Accumulate skills quickly

Strong risk-analytics teams use several roles to develop solutions and integrate them into business processes (Exhibit 4).

Recognizing that they might not have the time to build the whole arsenal of skills, leading banks have acquired companies, outsourced some analytical work, invested in fintechs, and entered into formal partnerships with analytic houses. JPMorgan Chase has partnered with OnDeck to lend to small businesses; Bank of America has committed $3 billion annually to fintech investment and joint innovation. Other leading banks have entered into partnerships with digital innovators to better understand customer behavior and risk profiles. Even when leading banks have acquired talent at scale in these ways, they still work to define roles and build skills in the risk-analytics team.

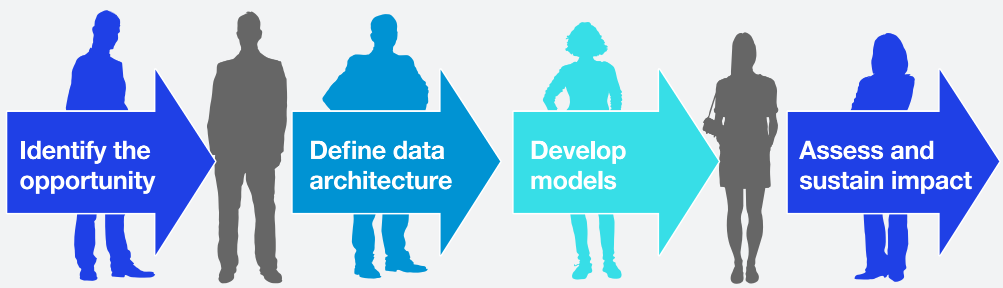
Fail often to succeed, iterating quickly

Speed is as important as completeness in realizing value from risk analytics. A winner-takes-all dynamic is emerging in the race to better serve customers. Banks, fintechs, and platform companies are getting better at locking in customers quickly with highly personalized and desirable offerings. The offerings are dependent on customer data, which get richer and deeper with every new development of risk-analytics capabilities.

Exhibit 4

Strong risk-analytics teams are using these roles to develop solutions and integrate them into business processes.

Structuring a strong risk-analytics team



Data engineers & data scientists: These roles are already common. What is new is that they encompass new techniques beyond traditional statistics and econometrics. Analytics teams now use such methods as graph theory to analyze supply-chain risk or machine-learning to develop highly sensitive early-warning systems.

Translators: This new role requires a keen business sense and an understanding of the rationale behind the models. It also requires an entrepreneurial spirit to promote risk analytics throughout the bank.

Business leaders and experts: Business leaders and experts are also involved in developing solutions, taking responsibility for embedding the risk model in current practices.

To reach and exceed the speed at which this race is moving, leading banks rely on quick, narrowly defined experiments designed to reveal the value (or the futility) of a particular hypothesis. When they succeed, they constitute a minimum viable product—something good enough to take to market, with the expectation that it will be soon improved. These experiments take weeks to conduct, rather than the more traditional months-long efforts commonly seen in risk-analytics functions (and that's not even considering the validation process). One form such experiments have taken are "hackathons"—coding sessions with analysts and others that have produced promising applications in compressed time frames.

Use model validation to drive relentless improvement

The banks that are developing a competitive edge through analytics constantly improve their current models, even as they build new ones. They make full use of their independent model-validation framework, moving beyond providing regulatory and statistical feedback on risk models every year to a more insightful and business-linked feedback loop. Validation departments can achieve this without losing their independence by changing from a mind-set of "examiners of models" to "guardians of model performance."

To introduce a degree of experimentation into model validation, leading banks incorporate business and model expertise into bursts of rapid development and testing, and accept that not all results will be as expected. In this way, the model benefits from a continual 360-degree review, rather than being buried in the risk-modeling team and understood only by the model owner. To be sure, as they do this work, banks must also respect regulatory constraints and explain to supervisors how they are utilizing advanced techniques. But leading institutions do not use regulatory oversight as an excuse not to move forward in an agile fashion. As shown by the multiple examples in this article, even large banks can make significant changes to improve outcomes and customer experience.

Getting started

We have outlined the reasons leading banks see considerable near-term promise in improved risk analytics, and the behaviors and principles that are distinguishing more successful players from the rest. This raises a logical question about what comes next: How can banks develop and execute a long-term bankwide risk-analytics strategy? While a full discussion is beyond the scope of this article, we see five immediate actions for the chief risk officer (CRO) to maximize the value of existing investments and prioritize new ones. These actions are all consistent with the six successful behaviors discussed above, but distilled into immediate high-payoff steps.

- Assess the current portfolio of risk-analytics projects, assets, and investments, and take a hard look at any that cannot answer the following questions satisfactorily:
  - Is the initiative business driven? Does it address one of the biggest business opportunities and define an analytics use case to deliver it? Or is the initiative a hammer looking for a nail?
  - Does the initiative have a clear plan for adoption and value capture? Or is it only a "model building" project?
  - Is the initiative structured to generate quick improvements as well as longer-term impact?
- Make an inventory of your talent, teams, and operating model for each initiative. Success requires multidisciplinary co-located teams of data engineers, data scientists, translators, and business experts. Prioritize actions to find the talent you need, rather than stretching the talent you have to the point of ineffectiveness.
- List your data and technology choke points—the weakest links in the system. Then determine the work-arounds you can develop to get high-priority initiatives moving (such as using external or alternative internal data or vendor solutions). Where no work-around is possible, ensure that precious resources do not lay idle waiting for resolution.

- Explain what you are doing to senior leaders, including business heads, the chief operating officer, and the chief investment officer. Work with them as needed to adjust priorities and redirect the program, but then proceed full steam ahead.

In our experience, risk leaders can take these steps quickly, given the right level of determination and focus. CROs should not hesitate to pull critical people into the exercise for a couple of weeks—it's typically a worthwhile investment that pays off in the redirection of a much larger body of work toward maximum impact.

Rajdeep Dash is a senior expert in McKinsey's London office, **Andreas Kremer** is a partner in the Berlin office, and **Luis Nario** and **Derek Waldron** are partners in the New York office.

About McKinsey & Company

McKinsey & Company is a global management consulting firm, deeply committed to helping institutions in the private, public, and social sectors achieve lasting success. For over eight decades, our primary objective has been to serve as our clients' most trusted external advisor. With consultants in 109 locations in over 60 countries, across industries and functions, we bring unparalleled expertise to clients anywhere in the world. We work closely with teams at all levels of an organization to shape winning strategies, mobilize for change, build capabilities, and drive successful execution. For more information, visit mckinsey.com.

January 2018

mckinsey.com
@McKAnalytics